UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL

ATUALIZAÇÃO DAS PREVISÕES DE CURTO PRAZO DE AFLUÊNCIAS AO SISTEMA HIDROELÉTRICO BRASILEIRO A PARTIR DA TÉCNICA DE PONDERAÇÃO BAYESIANA E DE PREVISÕES MENSAIS DE AFLUÊNCIA COM USO DE INFORMAÇÃO CLIMÁTICA

VINICIUS GROSSI DE OLIVEIRA

ORIENTADOR: CARLOS HENRIQUE RIBEIRO LIMA

DISSERTAÇÃO DE MESTRADO EM TECNOLOGIA AMBIENTAL E RECURSOS HÍDRICOS

PUBLICAÇÃO: PTARH.DM – 169/2015
BRASÍLIA/DF: MAIO – 2015

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL

ATUALIZAÇÃO DAS PREVISÕES DE CURTO PRAZO DE AFLUÊNCIAS AO SISTEMA HIDROELÉTRICO BRASILEIRO A PARTIR DA TÉCNICA DE PONDERAÇÃO BAYESIANA E DE PREVISÕES MENSAIS DE AFLUÊNCIA COM USO DE INFORMAÇÃO CLIMÁTICA

VINICIUS GROSSI DE OLIVEIRA

DISSERTAÇÃO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISÍTOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM TECNOLOGIA AMBIENTAL E RECURSOS HÍDRICOS.

APROVADA POR:

Prof. Carlos Henrique Ribeiro Lima, PhD (ENC-UnB)
(Orientador)

Prof. Dirceu Silveira Reis Junior, PhD (ENC-UnB)
(Examinador Interno)

Prof. Eduardo Sávio Passos Rodrigues Martins, PhD (UFC-FUNCEME)
(Examinador Externo)

BRASÍLIA/DF, MAIO DE 2015

FICHA CATALOGRÁFICA

OLIVEIRA, V.G.

ATUALIZAÇÃO DAS PREVISÕES DE CURTO PRAZO DE AFLUÊNCIAS AO SISTEMA HIDROELÉTRICO BRASILEIRO A PARTIR DA TÉCNICA DE PONDERAÇÃO BAYESIANA E DE PREVISÕES MENSAIS DE AFLUÊNCIA COM USO DE INFORMAÇÃO CLIMÁTICA.

xviii, 150p., 210 x 297 mm (ENC/FT/UnB, Mestre, Tecnologia Ambiental e Recursos Hídricos, 2015). Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Civil e Ambiental.

1. PREVISÃO DE VAZÕES	2. PONDERAÇÃO BAYESIANA DE MODELOS
3. INFORMAÇÕES CLIMÁTICAS	4. SISTEMA HIDROELÉTRICO BRASILEIRO
I. ENC/FT/UnB	II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

OLIVEIRA, V.G. (2015). *Atualização das previsões de curto prazo de afluências ao sistema hidroelétrico brasileiro a partir da técnica de ponderação bayesiana e de previsões mensais de afluência com uso de informação climática*. Dissertação de Mestrado em Tecnologia Ambiental e Recursos Hídricos, Publicação PTARH.DM-169/2015, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 150p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Vinicius Grossi de Oliveira

TÍTULO: Atualização das previsões de curto prazo de afluências ao sistema hidroelétrico brasileiro a partir da técnica de ponderação bayesiana e de previsões mensais de afluência com uso de informação climática.

GRAU: Mestre ANO: 2015

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Vinicius Grossi de Oliveira – vinicius.grossi@yahoo.com.br

Aos meus pais

À Paula

AGRADECIMENTOS

Ao meu pai, pelo exemplo de vida! Obrigado pelos seus ensinamentos, pela dedicação e perfeição em tudo o que faz.

À minha mãe, mulher e advogada batalhadora. Obrigado por ser minha maior incentivadora e a pessoa mais afetuosa do mundo.

À minha irmã, pelo carinho; e por mostrar que é possível.

Aos meus tios, tias, primos e primas. Adoro vocês!

À minha sogra, à minha cunhada e ao seu marido, meu amigo Fred. Obrigado pelo apoio e pela nova família. Uma honra ser o padrinho da pequena Carolina.

Aos amigos pra toda a vida, Rafael e Tulio. Não há palavras para descrever o carinho que tenho por vocês, meus irmãos!

À turma da Duque6, Ricardo, Junior, Andrey e Juliano, pela inoxidável amizade!

Aos amigos da Escola Federal de Engenharia de Itajubá, pelos fabulosos anos de convivência. Meus irmãos Tatá, Arthur, Claudião, Xerife, Osório, Ourinhos, Paulão, Bassê, Pig, Gozzy, Gus, Gutão, Guaxupé e todos os amigos da Confraria Mel da Lua. É muita gente boa por metro quadrado. Difícil citar todos, impossível esquecê-los!

Aos compadres, Alisson e Gisele, Galo e Michelle, e José Guilherme e Sinara.

Aos amigos itajubenses de Taubaté: Ourinhos, Negão, Belló, Pirangone, Felipão, Gerardo, suas esposas e filhos. Grande satisfação poder participar dessa grande família.

Aos amigos itajubenses de Brasília: Luciano, Renatão, Joaquim, Henrique, Gustavo, Hilton e respectivas famílias.

Aos amigos de Macapá: Marcos Branche, Janio, Ladeira e Rosielci. Obrigado por me propiciarem um dos melhores anos da minha vida!

Aos caríssimos Carlão, Baptistella, Altair, Julian, Bahia, Sr. Gunther, Jean Simal, Miltinho, Velho Bezerra, Vitorino, Benício, Wilson, Sidney, Maduro, Duarte, Emerson, Perrote, Gentil, todos os amigos barrageiros, onde estiverem, o meu muito obrigado!

Aos amigos Rafael Ribeiro, Cassio, Tarita, Ernandes, Zé Antonio, Gustavo, Guilherme, Kleverson, Paulinho, Osiane, Adriano, Krause, Rodrigo e Livia.

Aos amigos Ricardo, Flavio, Eduardo, William, Hugo, Leandro, Lucas, Nelson, José Henrique, José Antonio, Gentil, Nascimento, Vitor Hugo, Pedrão, Edesio, Zelismar, Sonia, Eliane, Amparo e demais colegas da área de logística.

Aos meus novos velhos amigos: Luís, Simone, Ana Lucia, Flavio e toda a turma.

Ao professor Carlos Henrique Ribeiro Lima, meu orientador, o meu muito obrigado por todas as intruções, transmissão de conhecimento, e pela disponibilidade oferecida durante a elaboração deste trabalho.

A todos os colegas do PTARH, sem exceção, agradeço pela experiência ímpar de convivência em uma atmosfera harmônica, transparente e de grande troca de experiências.

A todos os professores do PTARH e demais servidores, minha gratidão pela oportunidade de retornar ao ambiente acadêmico e sua efervescência de idéias. Não me restam dúvidas de que o ensino público de qualidade é o caminho para nosso país.

Agradeço aos membros da banca examinadora, os professores Dirceu Reis (UnB) e Eduardo Sávio Martins (UFC e FUNCEME), pela análise crítica do trabalho e pelas relevantes contribuições.

À ANEEL, em nome do Dr. Odenir, o meu muito obrigado pela oportunidade de cursar o mestrado, paralelamente às minhas atividades na SGH. Agradeço também aos amigos e colegas da ANEEL, com os quais tenho a honra de conviver profissional e pessoalmente há quase quatro anos.

Aos amigos Odair Manfroi, Jorge Caetano e Bruno Goulart. A conclusão do mestrado não seria possível sem o fundamental apoio de vocês.

Aos colegas do ONS, Vinicius Forain, Angela Ghirardi e Ana Paula Krisek, meu muito obrigado pela disponibilidade das informações e de seu precioso tempo.

Aos amigos da Alstom, onde iniciei minha carreira no setor elétrico, há quinze anos. Em especial, à turma da montagem externa e às equipes dos projetos que participei.

Aos amigos e colegas da Eletronorte e do Ministério de Minas e Energia, pelo convívio e aprendizado.

Em suma, o meu muito obrigado a todos que participaram da minha formação pessoal e profissional. Agradeceria a todos nominalmente, se aqui possível.

Por fim, um especial agradecimento à Paula, minha esposa, que acompanhou de perto todos os meus passos, do início ao fim do mestrado. Sem você, eu não teria conseguido. Obrigado pela paciência e pelas palavras de apoio nos momentos mais difíceis. Dedico esta conquista a você, meu amor!

RESUMO

O uso de sistemas eficientes de previsão de afluências de curto, médio e longo prazo permite otimizar a operação do conjunto de reservatórios hidroelétricos brasileiros, elevando o grau de segurança no fornecimento de energia elétrica e minimizando os custos operacionais. Os modelos atuais de previsão utilizados pelo Operador Nacional do Sistema Elétrico (ONS) tendem a ser limitados no horizonte de previsão e na modelagem da estrutura de dependência existente entre as diversas escalas temporais, reduzindo a qualidade das previsões e simulações de afluências.

Nesta dissertação, visando contribuir com a melhoria das previsões de vazão de curto prazo, são desenvolvidos e aperfeiçoados modelos estatísticos de previsão a partir de uma estrutura hierárquica, onde as vazões geradas para as menores escalas temporais (semanais) apresentam uma estrutura de dependência das vazões geradas para as maiores escalas (mensais).

As previsões mensais são obtidas por meio de um modelo periódico auto-regressivo exógeno (PARX), que utiliza informações climáticas de larga escala como covariáveis, de forma a aumentar os horizontes temporais de previsões das vazões e melhor representar a variabilidade espacial entre os postos fluviométricos. O modelo mensal PARX é aplicado na previsão de afluências aos vinte e oito principais reservatórios hidroelétricos de regularização do país. Previsões semanais para os mesmos reservatórios são obtidas a partir dos modelos utilizados pelo ONS e são atualizadas a partir da integração com as previsões mensais obtidas pelo modelo mensal PARX por meio da técnica de Ponderação Bayesiana de Modelos (BMA), que consiste em estimar pesos para os modelos semanal e mensal tendo como base o desempenho desses modelos, permitindo também obter uma estimativa das incertezas acerca dos mesmos. Assim, as previsões semanais atualizadas irão consistir de uma média ponderada das previsões obtidas pelos modelos semanal e mensal e o contexto Bayesiano de estimação dos pesos.

Os resultados obtidos por meio de validação cruzada apontam um ganho estatisticamente significante em termos de previsibilidade quando as previsões semanais disponibilizadas pelo ONS são integradas via BMA com as previsões mensais oriundas do modelo PARX, a partir da análise das quatro métricas de desempenho utilizadas – raiz do erro médio quadrático (RMSE), erro médio percentual absoluto (MAPE), indicador de Nash-Sutcliffe (NS) e distância multicritério (DM). Tendo como base os dados observados

e as previsões semanais utilizadas pelo ONS para o período entre janeiro de 2009 e setembro de 2014, observa-se um aumento médio na qualidade das previsões: 7,46% de ganho médio no indicador DM, para as previsões com uma semana de antecedência; 10,99%, para as previsões com duas semanas; 6,39%, nas previsões com três semanas; 10,08%, nas previsões com quatro semanas; 4,39%, nas previsões com cinco semanas de antecedência e 4,14%, com seis semanas. A partir desses resultados e em virtude da praticidade de implementação da técnica BMA nas atuais previsões semanais utilizadas pelo ONS, julga-se ser essa uma ferramenta eficaz e promissora para o aperfeiçoamento das previsões de vazão dos reservatórios do Sistema Interligado Nacional (SIN).

ABSTRACT

The use of efficient systems of short, medium and long term inflows forecast, allows optimizing the operation of the Brazilian interconnected hydroelectric reservoirs, raising the level of security of electricity supply and minimizing operating costs. The current forecast models used by the Brazilian Electric System Operator (ONS) tend to be limited in the forecast horizon and in the modeling of the dependence structure among the different time scales, thus reducing the quality of inflows forecasts and simulations.

In this dissertation, in order to contribute to the improvement of short-term flow forecasts, it has developed and improved statistical forecasting models based on a hierarchical structure, where the flows generated for smaller time scales (weekly) present a structure dependence on the flows generated for the larger scales (monthly).

The monthly forecasts are obtained from a periodic autoregressive exogenous model (PARX), which makes use of large-scale climate information as covariates in order to increase the time horizons of flow forecasts and better represent the spatial variability among gauging stations. The monthly PARX model is tested to inflows forecasts in the twenty eight major hydroelectric regularization reservoirs in the country. Weekly forecasts for the same reservoirs are obtained from the models used by ONS and updated from the integration with the monthly predictions obtained by the monthly model PARX through Bayesian Models Averaging technique (BMA), which is to estimate weights for the weekly and monthly models based on the performance of these models, also allowing to obtain an estimation of the uncertainty about them. Thus, updated weekly forecasts will consist of a weighted average of the forecasts obtained by weekly and monthly models and the Bayesian context of estimation of weights.

The results obtained by cross validation show a statistically significant gain in terms of predictability when the weekly forecasts provided by the ONS are integrated via BMA with monthly forecasts derived from the PARX model, based on the analysis of four performance metrics used - root mean square error (RMSE), mean absolute percentage error (MAPE), Nash-Sutcliffe indicator (NS) and multicriteria distance (DM). Based on the observed data and the weekly forecasts used by the ONS for the period between January 2009 and September 2014, there was an average increase in the quality of the forecasts: 7,46% average gain in the DM indicator, for predictions for one week lead time; 10,99% for two weeks forecasts lead time; 6,39%, for three weeks; 10,08%, for four weeks; 4,39%,

for five weeks forecasts lead time and 4,14%, for six weeks. From these results and because of the practicality of implementing the BMA technique on current weekly forecasts used by the ONS, is believed to be such an effective and promising tool for improving inflows forecasts to the National Interconnected Power System (SIN) reservoirs.

SUMÁRIO

LISTA DE TABELAS xiv

LISTA DE FIGURAS xv

LISTA DE NOMENCLATURAS E ABREVIAÇÕES xvii

1. INTRODUÇÃO 1

2. OBJETIVO 4

2.1. OBJETIVO GERAL 4

2.2. OBJETIVOS ESPECÍFICOS 4

3. REVISÃO DA LITERATURA 5

3.1. SÉRIES TEMPORAIS DE VAZÃO 5

3.1.1. Alguns modelos de séries temporais 7

3.1.2. Alguns resultados utilizando modelos de séries temporais 9

3.2. HISTÓRICO DA UTILIZAÇÃO DOS MODELOS DE PREVISÃO DE VAZÃO NO SETOR ELÉTRICO. 10

3.3. MODELOS DE PREVISÃO DE VAZÃO DO SETOR ELÉTRICO QUE NÃO UTILIZAM INFORMAÇÕES METEOROLÓGICAS 14

3.3.1. Modelos PREVIVAZ E PREVIVAZM 14

3.3.2. Modelo PREVIVAZH 17

3.3.3. Modelo CPINS 17

3.4. MODELOS DE PREVISÃO DE VAZÃO QUE UTILIZAM INFORMAÇÕES CLIMÁTICAS 18

3.4.1. Modelo de Previsão de Classes de Vazão (MPCV) 20

3.4.2. Modelo SMAP-MEL 20

3.4.3. Modelo SMAP 21

3.4.4. Modelo MGB 22

3.4.5. Modelo Fuzzy 23

3.5. VARIABILIDADE CLIMÁTICA E PADRÕES DE CHUVA E VAZÃO NO BRASIL 23

3.5.1. Padrões de teleconexão 24

3.5.2. El Niño/La Niña - Oscilação Sul 24

3.5.3. Mudanças nos padrões de chuva e vazão associadas com eventos de El Niño no Brasil..... 28

3.5.4. A zona de convergência do Atlântico Sul (ZCAS). 29

3.6. USO DE INFORMAÇÃO METEOROLÓGICA EM PREVISÕES DE VAZÃO 30

3.6.1. Alguns Resultados 31
3.7. O PROBLEMA DA INCERTEZA DOS MODELOS 34
3.8. A PONDERAÇÃO BAYESIANA DE MODELOS 36
4. REFERENCIAL TEÓRICO 41
4.1. O PARADIGMA BAYESIANO 41
4.1.1. Teorema de Bayes 41
4.2. PONDERAÇÃO BAYESIANA DE MODELOS 43
4.3. MODELOS AUTO REGRESSIVOS 44
5. METODOLOGIA 46
5.1. ESTRATÉGIA METODOLÓGICA 49
5.1.1. Modelo mensal 49
5.1.2. Modelo semanal 51
5.1.3. Ponderação dos modelos mensal e semanal 54
5.1.4. Métricas de qualidade adotadas 56
5.1.5. O teste estatístico de Diebold-Mariano 58
5.1.6. O pacote computacional *ensembleBMA* 60
6. RESULTADOS E DISCUSSÕES 62
6.1. RESULTADOS PARA AS MODELAGENS MENSAIS 62
6.1.1. Índices Climáticos 65
6.2. RESULTADOS PARA AS MODELAGENS SEMANAIS - DESEMPENHO DOS MODELOS DO ONS. 75
6.3. RESULTADOS PARA AS MODELAGENS SEMANAIS - PONDERAÇÃO BAYESIANA DOS MODELOS. 77
6.3.1. Resultado das previsões dos vinte e oito reservatórios em estudo. 78
6.3.2. Resultados espacializados dos indicadores de qualidade das previsões 84
6.3.3. Resultados localizados por bacias hidrográficas 90
6.3.4. O teste estatístico de Diebold-Mariano 99
7. CONCLUSÕES E RECOMENDAÇÕES 104
8. REFERÊNCIAS BIBLIOGRÁFICAS 108
APÊNDICES 120
APÊNDICE A - EXEMPLO SIMPLIFICADO DE APLICAÇÃO DA TÉCNICA BMA (EXTRAÍDO DE GIBBONS *ET AL*., 2008) 121
APÊNDICE B - O ALGORITMO *EXPECTATION-MAXIMIZATION* 125

APÊNDICE C - DESEMPENHO DAS PREVISÕES DOS MODELOS UTILIZADOS PELO ONS EM RELAÇÃO ÀS VAZÕES OBSERVADAS (DISTÂNCIA MULTICRITÉRIO ANUAL ENTRE 2009 E 2013). 129

APÊNDICE D – DESEMPENHO DAS SIMULAÇÕES POR RESERVATÓRIO 132

APÊNDICE E - INDICADORES DE QUALIDADE MÉDIOS (COMPARATIVO DOS RESULTADOS BMA EM RELAÇÃO AO DESEMPENHO DOS MODELOS UTILIZADOS PELO ONS ENTRE JAN-2009 E SET-2014). 147

LISTA DE TABELAS

Tabela 3.1 - Modelos de previsão utilizados no setor elétrico até 2013..... 13
Tabela 3.2 - Modelos implementados no PREVIVAZ..... 16
Tabela 3.3 - Modelos implementados no PREVIVAZM..... 16
Tabela 3.4 - Desvios médios absolutos de previsão de vazões dos modelos PREVIVAZ (entre 1997 e 2004) e PREVIVAZH (entre 2000 e 2001)..... 19
Tabela 5.1 - Exemplo das informações oriundas dos relatórios do PMO..... 52
Tabela 5.2 – Esquema de reorganização dos dados de acordo com a antecedência semanal 53
Tabela 5.3 – Dados da Tabela 5.1, ordenados de acordo com a antecedência das previsões semanais. 53
Tabela 5.4 – Visão esquemática da matriz de dados primários para a composição BMA. 54
Tabela 6.1 - Identificação dos 28 reservatórios que fazem o escopo deste estudo, sua localização (rio), a área de drenagem aproximada e o(s) modelo(s) de previsões semanais utilizados pelo ONS. 63
Tabela 6.2 – Valores médios do indicador NS para cada período de calibração ilustrado na Figura 6.2..... 65
Tabela 6.3 - Número de amostras no período de treinamento e de validação. 77
Tabela 6.4 - Resultados para o indicador DM na UHE Sobradinho entre 2009 e 2013. 99
Tabela 6.5 - Resultados do teste de Diebold-Mariano para todos os vinte e oito reservatórios (casos de rejeição da hipótese nula, ao nível de significância de 10%). 101
Tabela 6.6 - Resumo de alguns dos resultados por reservatório. 103

LISTA DE FIGURAS

Figura 3.1 - Esquema do Modelo SMAP aplicado para a bacia do Rio Grande com 4 reservatórios. 22
Figura 3.2 - Oceano Pacífico em condições normais 26
Figura 3.3 - Oceano Pacífico em condições de El Niño..... 26
Figura 3.4 - Oceano Pacífico em condições de La Niña 27
Figura 3.5 - Regiões do ENOS..... 27
Figura 3.6 - Típicos padrões de chuva e temperatura associados com a ocorrência de El Niño. 29
Figura 5.1 - Comparação entre vazão observada e a estimada na semana corrente – bacia do rio Paraná – na UHE Itaipu..... 46
Figura 5.2 - Comparação entre vazão observada e prevista com uma (1) semana de antecedência – bacia do rio Paraná – na UHE Itaipu. 47
Figura 5.3 - Comparação entre vazão observada e prevista com duas (2) semanas de antecedência – bacia do rio Paraná – na UHE Itaipu. 47
Figura 5.4 - Comparação entre vazão observada e prevista com três (3) semanas de antecedência – bacia do rio Paraná – na UHE Itaipu. 48
Figura 5.5 - Representação gráfica da distância multicritério.. 58
Figura 6.1 - Localização dos sítios hidroelétricos utilizados na modelagem. 62
Figura 6.2 – Comparativo do indicador NS entre diversos períodos de calibração para o modelo PARX mensal. 65
Figura 6.3 - Localização espacial dos índices climáticos baseados na SST (retângulos com linha sólida, sendo que o mais à esquerda se refere ao índice NINO3) e no vento zonal de baixa pressão (retângulo com linha tracejada).. 66
Figura 6.4 - Correlações entre as vazões do mês de março e o índice NINO3, com *lags* de um e dois meses.. 67
Figura 6.5 - Correlações entre as vazões do mês de setembro e o índice NINO3, com *lags* de um e dois meses. 68
Figura 6.6 - Valores médios anuais das correlações entre as vazões e o índice NINO3, com *lags* de um e dois meses. 68
Figura 6.7 - Correlações entre as vazões do mês de março e o índice SST, com *lags* de um e dois meses.. 69
Figura 6.8 - Correlações entre as vazões do mês de setembro e o índice SST, com *lags* de um e dois meses.. 70
Figura 6.9 - Valores médios anuais das correlações entre as vazões e o índice SST, com *lags* de um e dois meses..... 70
Figura 6.10 - Valores médios anuais das correlações entre as vazões e o índice de vento zonal (U1), com *lags* de um e dois meses. 71
Figura 6.11 - Funções periódicas de auto-correlação com *lags* de um e dois meses.. 72
Figura 6.12 - Distribuição espacial dos índices NS para o modelo PARX com *lags* de um a dois meses, no mês de março.. 73
Figura 6.13 - Assim como na Figura 6.12, mas para o mês de setembro. 73
Figura 6.14 - Comparativo entre os indicadores globais NS dos modelos PARX e PAR com *lag* de um mês. 74
Figura 6.15 - Assim como na Figura 6.14, mas para o *lag* de dois meses. 74

Figura 6.16 - Resultados do indicador distância multicritério para os modelos utilizados pelo ONS com antecedência de 1 semana (média do período entre 2009 e 2013). 75
Figura 6.17 – Assim como na Figura 6.16, mas para a antecedências de 2 semanas. 76
Figura 6.18 - Assim como na Figura 6.16, mas para a antecedências de 3 semanas. 76
Figura 6.19 - Distribuição dos pesos Bayesianos para a UHE Emborcação para cada antecedência semanal (período completo).. 78
Figura 6.20 - Distribuição espacial dos pesos dos modelos de previsão mensal para os 28 reservatórios analisados, nas previsões com antecedências de uma (esquerda) e duas (direita) semanas. 80
Figura 6.21 – Assim como na Figura 6.20, mas para previsões realizadas com três (esquerda) e quatro (direita) semanas de antecedência. 80
Figura 6.22 - Indicadores de desempenho para previsões realizadas com o modelo mensal de uma a seis semanas de antecedência para as afluências à UHE Emborcação. 81
Figura 6.23 - Valores para o indicador DM com uma a seis semanas de antecedência na UHE Emborcação 82
Figura 6.24 – Assim como na Figura 6.20, mas para previsões realizadas com cinco (esquerda) e seis (direita) semanas de antecedência. 83
Figura 6.25 - Indicadores com uma a seis semanas de antecedência na UHE Tucuruí para o modelo mensal em relação aos valores verificados semanais (período completo). 83
Figura 6.26 - Resultados das diferenças entre o indicador RMSE das previsões do modelo semanal e das previsões BMA para as seis semanas de antecedência 85
Figura 6.27 - Resultados das diferenças entre o indicador MAPE das previsões dos modelos semanais e das previsões BMA para as seis semanas de antecedência. 86
Figura 6.28 - Resultados das diferenças entre o indicador NS das previsões semanais e das previsões BMA para as seis semanas de antecedência. 88
Figura 6.29 - Resultados das diferenças entre o indicador DM das previsões semanais realizadas pelo ONS e das previsões BMA de uma a seis semanas de antecedência. 89
Figura 6.30 - Esquema topológico das usinas das bacias do Paranaíba e Grande 90
Figura 6.31 - Esquema topológico das usinas do Tietê-Paranapanema. 93
Figura 6.32 - Esquema topológico das usinas do rio Iguaçu. 95
Figura 6.33 - Esquema topológico das usinas da bacia dos rios Uruguai e Jacuí. 97
Figura 6.34 - Esquema topológico das usinas das bacias dos rios São Francisco e Tocantins. 98
Figura 6.35 - Resultados do teste estatístico de Diebold-Mariano para a UHE Tucuruí, sob a hipótese nula de que os erros das previsões de vazões dos dois modelos são iguais.. 100

LISTA DE NOMENCLATURAS E ABREVIAÇÕES

AIC - Critério de Informação de Akaike

AMO - *Atlantic Multi-decadal Oscillation*

ANEEL - Agência Nacional de Energia Elétrica

ANN - Rede Neural Artificial

AR - Modelo auto-regressivo

ARMA - Modelo auto-regressivo de médias móveis

BIC - Critério de Informação Bayesiana

BMA - *Bayesian Model Averaging*

CEPEL - Centro de Pesquisas de Energia Elétrica, da Eletrobrás.

CPINS - Cálculo e Previsão de Vazões Incrementais e Naturais a Sobradinho

CPTEC - Centro de Previsão de Tempo e Estudos Climáticos

DIANA - Modelo de geração de séries sintéticas de vazões diárias (CEPEL)

DM - indicador de qualidade distância multicritério

Eletrobras - Centrais Elétricas Brasileiras S.A.

ENOS - El Niño Oscilação Sul

EM - algoritmo *expectation maximization*

EPSs - *Ensemble Prevision Systems*

FCTH - Fundação Centro Tecnológico de Hidráulica

FDP - Função densidade de probabilidade

GA - Algoritmo Genético

GCM - Modelo Climático de Circulação Geral

ICA - Análise de Componentes Independentes

IRI - *International Research Institute for Climate and Society*

INPE - Instituto Nacional de Pesquisas Espaciais.

IPH - Instituto de Pesquisas Tecnológicas

LDA - Análise do Discriminante Linear

MAPE - indicador de qualidade do erro médio percentual absoluto

MCMC - Monte Carlo *Markov Chain*

MEL - Modelagem estocástica linear

MGB - Modelo de Grandes Bacias

MPCV - Modelo de Previsão de Classes de Vazão

NAO - Oscilação do Atlântico Norte

NCAR - *National Center for Atmospheric Research*

NCEP - *National Centers for Environmental Prediction*

NEURO3M - Modelo de Redes Neurais aplicado ao trecho a montante da UHE Três Marias

NeuroSF - Modelo de Redes Neurais aplicado ao trecho entre UHE Três Marias e a UHE Sobradinho

NOAA - *National Oceanic and Atmospheric Administration*

NS - indicador de qualidade Nash-Sutcliffe

OLAP - *On-Line Analytical Procedure*

ONS - Operador Nacional do Sistema Elétrico

PAR - Modelo periódico auto-regressivo

PARMA - Modelo periódico Auto-regressivo de médias móveis

PARX - Modelo periódico auto-regressivo exógeno

PCA - Análise de Componentes Principais

PMO - Programa Mensal de Operação

PDO - *Pacific Decadal Oscillation Index*

PMP - *Posterior Model Probability*

RMSE - indicador de qualidade da raiz do erro médio quadrático

PREVIVAZ, PREVIVAZM, PREVIVAZH - família de modelos de previsão de vazão desenvolvidos pelo CEPEL

SARIMA - Modelo Sazonal Auto-regressivo Integrado de Média Móvel

SIN - Sistema Interligado Nacional

SMAP - *Soil Moisture Accounting Procedure*

SOI - *Southern Oscillation Index*

SSARR - *Streamflow Synthesis and Reservoir Regulation*

SST - temperatura da superfície do mar (*sea surface temperature*)

SVD - Decomposição do Valor Singular

U1 - índice de anomalia do vento zonal de baixa pressão

UHE - Usina Hidroelétrica

ZCAS - Zona de Convergência do Atlântico Sul

ZCIT - Zona de Convergência do Intertropical

1. INTRODUÇÃO

O conhecimento prévio acerca das vazões naturais afluentes aos aproveitamentos hidroelétricos é insumo fundamental para o planejamento e programação da operação de um parque gerador de energia elétrica, sobretudo do Sistema Interligado Nacional (SIN), cuja predominância da geração é originada de fonte hidroelétrica.

Devido à irregularidade das vazões e à necessidade de manter a continuidade do fornecimento de energia elétrica, o sistema de geração de energia elétrica brasileiro, além de contar com usinas termoelétricas de complementação (o que implica em custos com combustíveis), conta principalmente com um conjunto de reservatórios de acumulação, cuja principal função é a regularização dos regimes fluviais, através do armazenamento de água nos períodos de maiores afluências naturais para posterior utilização durante períodos mais secos, quando são escassas as vazões.

A coordenação conjunta da operação do sistema de reservatórios das hidrolétricas e do sistema de usinas alimentadas pelas demais fontes permite o melhor aproveitamento das vazões naturais, reduzindo o desperdício de água. O uso ineficaz da água na operação do sistema de reservatórios poderá incorrer em uma série de desperdícios, tais como vertimentos desnecessários, evaporação dos reservatórios e operação das centrais hidroelétricas fora de sua faixa de maior eficiênica. Quanto melhores os sistemas de previsão de vazões, maiores as possibilidades de alocação correta da água, reduzindo o uso de combustíveis fósseis para complementar o fornecimento de energia e, assim, minimizar os custos de produção.

O planejamento e a programação da operação do SIN são feitos pelo Operador Nacional do Sistema Elétrico (ONS). A maioria dos modelos atuais de previsões de vazão adotados pelo Operador pertence à família de modelos estatísticos periódicos auto-regressivos (PAR), construídos individualmente para cada reservatório.

A construção de modelos autoregressivos (AR) adota como premissa básica a independência mútua das séries de afluência, uma vez conhecidas as co-variáveis utilizadas nos modelos. Embora esta suposição possa garantir que em previsões de curto prazo (em geral, de até doze dias à frente) seja mantida a variabilidade espacial observada nos dados históricos, a persistência da série de vazões e o uso de informações passadas tende a explicar pouco a variância de vazões de longo prazo, o que pode acarretar na violação dessa premissa (Lima, 2014).

Desde o período antecedente à criação do ONS até meados de 2008, definiu-se a utilização do programa PREVIVAZ, modelo estocástico univariado desenvolvido pelo Centro de Pesquisas de Energia Elétrica (CEPEL) da Eletrobras, como o modelo oficial de previsões de vazões das usinas do SIN.

Ainda em 2006, foi autorizada pela Agência Nacional de Energia Elétrica (ANEEL), órgão regulador do setor elétrico brasileiro, a utilização de um modelo para a bacia do rio Iguaçu, complementar ao PREVIVAZ, o Modelo de Previsão de Classes de Vazão (MPCV). Foi a primeira experiência da utilização de modelos que utilizam informações de precipitação observada e prevista.

Nesse sentido, Collischonn *et al*. (2007) mostram que a incorporação de previsões quantitativas de chuva na metodologia de previsão de vazão pode resultar na diminuição de vertimentos desnecessários, na redução do risco de inundações e na melhor operação do sistema. Os autores alegam, entretanto, que a precipitação é provavelmente a variável mais difícil de ser prevista utilizando os modelos numéricos de previsão de tempo e clima.

A partir de junho de 2008, o ONS passou a utilizar em algumas bacias modelos de previsão de vazões que incorporam informação meteorológica por meio da previsão de precipitação. A maior parte dos modelos nessa categoria é do tipo chuva-vazão e são adotados apenas em algumas bacias hidrográficas. Devido às limitações nas previsões de chuvas, o uso de tais modelos fica restrito a previsões de vazão de curto prazo. Ademais, as previsões de vazão realizadas nessa escala temporal são independentes daquelas geradas para as maiores escalas (ex. mensal e anual), onde a persistência é a informação relevante utilizada nos modelos utilizados. Nesse caso, é possível que a informação quantitativa das previsões de longo prazo seja pouco útil nas previsões de curto prazo.

Como alternativa e complementação aos modelos chuva-vazão acoplados a modelos numéricos de previsão de chuva, aparecem os modelos estatísticos que ligam diretamente as variáveis hidrológicas de interesse (vazão, por exemplo) a forçantes externas (ou seja, preditores climáticos), que tendem a ser de larga escala e de evolução mais lenta (por exemplo, temperatura da superfície do mar), propiciando assim previsões de longo prazo que tendem a possuir uma menor variância e tendenciosidade quando comparadas com aquelas obtidas por meio de modelos numéricos ou auto-regressivos. Exemplos de aplicações com ganhos na escala mensal podem ser vistos em Uvo e Graham (1998), Kelman *et al.* (2000), Souza Filho e Lall (2003), Araghinejad *et al.* (2006), Maity e

Kumar (2008) e Lima e Lall (2010). Para aplicações em maiores escalas temporais, sugere-se consultar os trabalhos de Tippett *et al.* (2004) e Tootle *et al.* (2008),

No intuito de contribuir com a melhoria das previsões de vazão de curto prazo, propõe-se neste trabalho integrar as previsões semanais atualmente realizadas pelo ONS com previsões mensais realizadas a partir de um modelo periódico auto-regressivo (PARX) que considera a informação climática e com isso a variabilidade espaço-temporal das séries históricas de afluências. A integração das previsões multi-escalas é realizada pela técnica de ponderação bayesiana de modelos (BMA), onde os pesos e respectivas incertezas atribuídos a cada modelo são estimados num contexto Bayesiano a partir do desempenho individual observado em um determinado conjunto de dados. O ganho em termos de qualidade nas previsões de vazões de curto prazo é quantificado no presente trabalho a partir da aplicação do modelo proposto em dados de vazão natural afluente aos vinte e oito principais reservatórios hidroelétricos do país e nas respectivas previsões semanais disponibilizadas pelo ONS no Programa Mensal de Operação (PMO).

Após a apresentação inicial, este trabalho está organizado em oito capítulos. O capítulo 2 trata dos objetivos desta dissertação. No capítulo 3 é apresentada a revisão da literatura, que traz as informações referentes às experiências com previsões de vazão e o detalhamento dos modelos de previsão utilizados no setor elétrico. Na sequência, apresentam-se alguns conceitos de climatologia e exemplos da utilização de preditores climáticos para melhoria das previsões de vazão, bem como a metodologia de ponderação bayesiana de modelos. O referencial teórico encontra-se no Capítulo 4 enquanto a metodologia adotada para a modelagem mensal e semanal e para a ponderação dos modelos é apresentada no capítulo 5. Os resultados são apresentados e discutidos no capítulo 6 e as conclusões e recomendações seguem no capítulo 7. No último capítulo consta a bibliografia utilizada.

2. OBJETIVO

2.1. OBJETIVO GERAL

Aperfeiçoar as previsões de afluências de curto prazo aos reservatórios do sistema hidroelétrico brasileiro a partir da integração com modelos estatísticos de previsão na escala mensal que consideram informação climática.

2.2. OBJETIVOS ESPECÍFICOS

1) Desenvolver e utilizar um modelo estatístico de previsões mensais de afluências que considere informação climática de larga escala;
2) Desenvolver uma metodologia que permita o acoplamento das previsões em diversas escalas temporais (semanal, mensal, etc.);
3) Avaliar o ganho da modelagem proposta sobre os modelos semanais atualmente em uso, por meio de comparação entre os indicadores de qualidade das previsões dos modelos (proposto e em vigor).

3. REVISÃO DA LITERATURA

O capítulo está dividido em tópicos, onde serão apresentados, de forma resumida, alguns estudos realizados em modelagem de séries temporais de vazão; um histórico dos métodos de previsão de vazões utilizados no setor elétrico brasileiro; alguns conceitos sobre variabilidade climática e sua influência sobre o Brasil e de que forma essas informações climáticas podem ser associadas às previsões de vazões; e, por último, detalhes acerca da incerteza de modelos e da técnica de ponderação bayesiana de modelos.

3.1. SÉRIES TEMPORAIS DE VAZÃO

A análise e a modelagem de séries temporais estocásticas são importantes ferramentas em Hidrologia e Recursos Hídricos. São usadas na construção de modelos matemáticos, com o objetivo de gerar registros hidrológicos sintéticos, determinar a probabilidade de eventos extremos, investigar tendências e mudanças em registros hidrológicos, preencher e estender dados com falhas, entre outros (Tesfaye *et al*., 2006).

Um objetivo importante em hidrologia estocástica é gerar chuvas sintéticas e/ou sequências de vazões que têm estatísticas e estruturas de dependência semelhantes às do registro histórico. Essas sequências representam chuvas futuras plausíveis e/ou cenários de vazões que podem ser usados como entradas de uma gama de aplicações, tais como a concepção e operação de reservatórios, sistemas de irrigação e sistemas hidroelétricos (Westra *et al*., 2007).

A geração de séries sintéticas de vazão em rios pode ser útil para determinar as dimensões de obras hidráulicas, avaliar o risco para abastecimento de água urbana e irrigação, otimizar a operação de sistemas de reservatórios, determinar o risco de falhas de confiabilidade em sistemas hidroelétricos, planejar a expansão da capacidade dos sistemas de abastecimento de água, entre outros (Salas, 1993).

Mesmo as pequenas melhorias no funcionamento dos grandes sistemas de reservatórios poderão resultar em economia de vários milhões de dólares por ano (Brochu, 1978; apud Noakes *et al*., 1985).

Modelos estocásticos de vazões são frequentemente utilizados em estudos de simulação para avaliar o desempenho desejável de sistemas de recursos hídricos no futuro (Stedinger e Taylor, 1982). Modelos estocásticos são frequentemente ajustados aos dados

históricos, a fim de produzir cenários de vazão. Estes cenários são utilizados como dados de entrada para simulação/otimização de modelos de suporte às decisões operacionais em sistemas de recursos hídricos. Os cenários de vazão são amostras retiradas de distribuições de probabilidade condicionadas às informações disponíveis, tais como informações recentes de vazão (Kelman *et al*., 2000).

Os primeiros esforços para gerar séries temporais hidrológicas vêm da década de 1960. Embora originalmente motivado por possíveis aplicações em finanças, o trabalho em análise de séries temporais de Box e Jenkins (1970) obteve um impacto profundo sobre os pesquisadores em hidrologia estocástica nas últimas três décadas (Ilich e Despotovic, 2008).

Para Tesfaye *et al.* (2006), as características estatísticas da série hidrológica são fatores de decisão importantes na escolha do tipo de modelo. Na maioria dos casos conhecidos na natureza, vazões em rios têm comportamento periódico significativo na média, desvio padrão e assimetria. Em adição a estas periodicidades, elas apresentam uma estrutura de correlação temporal que pode ser constante ou periódica. Tal dependência ou autocorrelação nas séries de vazão geralmente surgem a partir do efeito do armazenamento, seja ele superficial, do solo ou subterrâneo, que faz com que a água permaneça no sistema através de períodos de tempo subsequentes.

Não há atualmente uma metodologia universalmente aceita nem programa de computador disponível que tenha aceitação generalizada entre os hidrólogos (Ilich e Despotovic, 2008). Os motivos podem ser muitos, e eles são devidos em parte à natureza complexa dos processos hidrológicos, incluindo as descontinuidades dos dados em regiões secas, a controvérsia a respeito do fenômeno de Hurst (Hurst 1957), ou a incerteza proposta por alguns pesquisadores a respeito da suposição de que os fenômenos hidrológicos são, de fato, estacionários no longo prazo (Sveinsson e Salas 2003).

Como Srinivas e Srinivasan (2005) apontam, apesar dos inúmeros relatos sobre os esforços de modelagem em hidrologia estocástica, nenhum ganhou aceitação universal (Ilich e Despotovic, 2008).

Stedinger e Taylor (1982) apontam as seguintes etapas básicas que podem envolver o desenvolvimento e uso de um modelo estocástico de vazões:

1º) Obter registros de vazão e outras informações;

2º) Selecionar modelos para descrever as distribuições de probabilidade marginais de vazões em diferentes estações do ano e estimar os parâmetros dos modelos;

3º) Escolher um modelo adequado à dependência espacial e temporal das vazões;

4º) Verificar se a implementação computacional do modelo está de acordo com a especificada;

5º) Validar o modelo para a simulação do sistema hidrológico desejado;

6º) Utilizar o modelo.

Stedinger e Taylor (1982) ressaltam a distinção entre as etapas de verificação e validação do modelo, citando trabalhos anteriores (Schlesinger *et al*., 1979; Mihram 1972). A verificação do modelo deve demonstrar que um modelo idealizado foi implementado corretamente; já a validação do modelo é uma tarefa adicional e a mais difícil, que compara os resultados da simulação com os dados do sistema real para demonstrar que o modelo é uma descrição adequada do mundo real para a investigação pretendida.

As séries temporais de vazão exibem uma estrutura de autocorrelação que depende não só na defasagem de tempo entre as observações, mas também do período do ano (Moss e Bryson,1974; apud Noakes *et al*., 1985).

A geração de chuvas sintéticas multivariadas e/ou séries temporais de vazões que simulem com precisão tanto a dependência espacial quanto temporal da série multivariada original é um problema desafiador em hidrologia e frequentemente exige a estimativa simultânea de um grande número de parâmetros do modelo ou hipóteses simplificadoras significativas sobre a estrutura do modelo (Westra *et al*., 2007).

3.1.1. Alguns modelos de séries temporais

As contribuições de Box- Jenkins (1970) foram o marco inicial da modelagem estocástica, coincidente com a introdução dos modelos dos tipos auto-regressivos (AR) e auto-regressivos de médias móveis (ARMA).

Os modelos auto-regressivos buscam preservar a média, o desvio padrão e a estrutura de correlação das séries originais, sob a suposição de que os dados são

normalmente distribuídos, o que frequentemente requer que os dados sejam transformados antes da análise. Alternativamente, Westra *et al.* (2007) elencam algumas abordagens não paramétricas existentes que não requerem pressupostos anteriores sobre a natureza da distribuição de probabilidade, como por exemplo, os estudos de Lall *et al*, 1996; Lall e Sharma, 1996; Sharma, 2000b; Sharma e Neill, 2002; Sharma *et al.* 1997.

Westra *et al.* (2007) registram haver grande volume na literatura em modelagem de séries temporais hidrológicas univariadas, da qual a classe de modelos AR e ARMA são as mais comuns. São citados os estudos de Box *et al*., 1994; Bras e Rodrigues-Iturbe, 1985; Loucks *et al*., 1981 e Salas, 1992, em particular, para as séries temporais em escalas mensais ou acima.

Há um certo número de variantes do método de Box e Jenkins em hidrologia estocástica que promove os modelos ARMA em várias formas, mais freqüentemente em combinação com modelos de desagregação, que exigem que a série anual seja gerada previamente, assegurando que as estatísticas anuais sejam preservadas e, em seguida, divididas em intervalos de tempo sazonais (tipicamente mensais) usando vários algoritmos de desagregação (Ilich e Despotovic, 2008).

A modelagem estocástica das vazões sazonais provou ser difícil devido à natureza periódica dos parâmetros estocásticos juntamente com as reconhecidas dificuldades de modelagem do comportamento auto-regressivo de ordens superiores, restringindo, assim, a maioria dos esforços até hoje, principalmente para a modelagem mensal auto-regressiva - AR (1) (Ilich e Despotovic, 2008).

Existe certo número de métodos multivariados, como uma extensão multivariada para o conjunto de modelos ARMA, que buscam manter a estrutura de covariância da série temporal observada. O artigo de Westra *et al.* (2007) cita algumas referências (Pegram e James, 1972; Salas, 1992; Wilks, 1995).

No tocante aos casos em que se exige a modelagem multivariada, além de simular a dependência temporal, é também necessário se concentrar em manter a dependência espacial. O problema com estas abordagens é que normalmente é necessário estimar um grande número de parâmetros a partir de um número limitado de dados, o que pode tornar a abordagem consideravelmente mais difícil de aplicar em relação aos métodos univariados devido a grande incerteza nos parâmetros estimados (Westra *et al*., 2007).

A dificuldade principal com o modelo ARMA multivariado é que os parâmetros são frequentemente estimados usando o método de momentos, que considera apenas a

estrutura de covariância da série temporal multivariada, resultando na manutenção das estatísticas de momento de ordem um e dois (Westra *et al*., 2007). Assim, ao se avaliar um hipotético sistema interligado por cinco reservatórios e séries mensais com cinquenta anos de dados de vazão disponíveis, para cada previsão mensal o número de parâmetros a serem estimados chegaria a cento e cinquenta, o que demandaria um grande esforço computacional.

Tesfaye *et al.* (2006) sugererm a utilização de uma classe importante de modelos periódicos: os modelos periódicos auto-regressivos de média móvel (PARMA), que permitem remover a estrutura de correlação periódica dos modelos. Segundo os autores, os modelos PARMA representam explicitamente as flutuações sazonais na média, no desvio padrão e na autocorrelação das vazões, resultando em simulações mais confiáveis nas vazões naturais dos rios. A estimativa de parâmetros em modelos PARMA, porém, é mais difícil do que nos modelos ARMA estacionários, por causa do maior número de parâmetros a serem estimados.

3.1.2. Alguns resultados utilizando modelos de séries temporais

Stedinger e Taylor (1982) utilizaram dados de vazões da bacia do alto Delaware, nos Estados Unidos, para a verificação e validação de um modelo de vazões e para comparar as características das vazões geradas para uma vasta gama de modelos de vazões.

Noakes *et al.* (1985) descreveram uma família de modelos periódicos auto-regressivos (PAR) projetados especificamente para explicar as variações nas vazões mensais de trinta rios na América do Norte e do Sul, comparando o desempenho dos modelos PAR com o de modelos sazonais auto-regressivos integrados de média móvel (SARIMA) e modelos ARMA. Os resultados indicaram que o modelo PAR, usando a função de autocorrelação parcial, obteve as melhores previsões.

Oliveira *et al.* (1988) apresentaram uma metodologia de desagregação para a geração de vazões mensais afluentea a nove hidroelétricas brasileiras, alternativa à proposta por Pereira *et al.* (1984), com correlação cruzada correlacionada em vários locais ao mesmo tempo usando modelos de desagregação para as vazões em cada local. A abordagem proposta preserva as características univariadas do modelo de desagregação de vazões em cada estação, evitando assim a complexidade introduzida pela utilização dos grandes modelos multivariados.

Tesfaye *et al.* (2006) analisaram técnicas de simulação baseadas em um modelo PARMA para capturar as variações sazonais nas estatísticas de dados mensais de vazão no Rio Fraser, em British Columbia, Estados Unidos. Eles identificaram que os resíduos do modelo se aproximam da sequência fundamental de ruídos da qual o processo PARMA é construído. Por isso, é necessário estimar a distribuição estatística destas variáveis aleatórias, a fim de simulá-las com precisão, evitando resultados de simulação distorcidos, particularmente em termos de valores extremos, que são importantes para a análise de cheias e de estiagens.

Westra *et al.* (2007) propuseram uma abordagem em duas etapas para gerar séries temporais sintéticas multivariadas em períodos mensais ou maiores. Primeiro, transformando os dados em um conjunto de séries temporais univariadas estatisticamente independentes e, em seguida, aplicando um modelo de séries temporais univariadas para os dados transformados. Foi utilizada uma técnica de transformação conhecida como Análise de Componentes Independentes (ICA), que usa uma aproximação mútua das informações para maximizar a independência entre as séries transformadas. Os resultados do método ICA foram comparados com a técnica de Análise de Componentes Principais (PCA), que apenas remove a covariância (ou correlação espacial) das séries temporais multivariadas, sem garantir, necessariamente, a independência completa. Ambos os métodos foram testados em dados de afluências mensais aos reservatórios da Colômbia. Eles mostraram que a discrepância entre os dados gerados sinteticamente e os dados originais é reduzida em 25% quando se utiliza ICA comparada com PCA para a distribuição conjunta total e de 28% quando se considera as densidades marginais isoladas. Estes resultados sugerem que benefícios significativos poderão ocorrer ao maximizar-se a independência estatística no desenvolvimento de modelos de geração sintética de séries temporais multivariadas, em vez de simplesmente remover a correlação.

3.2. HISTÓRICO DA UTILIZAÇÃO DOS MODELOS DE PREVISÃO DE VAZÃO NO SETOR ELÉTRICO

O sistema hidroelétrico brasileiro é talvez o único em que as seqüências de vazões estocásticas multivariadas têm sido usadas rotineiramente para testar políticas operacionais e para auxiliar nas decisões de expansão da capacidade (Lepecki e Kelman, 1985; Terry *et al*, 1986; Pereira, 1985; apud Oliveira *et al*., 1988)

O planejamento e programação da operação do SIN são feitos atualmente pelo ONS, que possui dentre suas atribuições principais a previsão de vazões semanais para o PMO. Esta previsão de vazões semanal é elaborada para todas as semanas do mês do PMO a partir da execução de seus modelos.

Desde o período antecedente à criação do ONS, até meados de 2008, definiu-se a utilização do programa PREVIVAZ, modelo estocástico univariado desenvolvido pelo CEPEL, como o modelo oficial de previsões de vazões das usinas do SIN. Exceção feita à previsão de vazões na bacia do rio São Francisco, no trecho incremental entre a UHE Três Marias e a UHE Sobradinho, em que o Modelo de Previsão de Vazões Incrementais e Naturais a Sobradinho (CPINS) foi estabelecido como o modelo de previsões.

No decorrer do ano de 2005, o ONS, no âmbito do seu plano de ação, contratou o desenvolvimento de diferentes tecnologias para a previsão de vazões diárias, para um horizonte de 12 dias à frente, incorporando, além das vazões naturais passadas, as informações da precipitação observada e prevista. Para esse projeto foram escolhidas três bacias piloto: bacia do rio Iguaçu, do rio Paraná (Trecho entre a UHE Rosana, a UHE Porto Primavera e a UHE Itaipu) e do rio Paranaíba (trecho entre a UHE Itumbiara e a UHE São Simão).

Para o desenvolvimento das metodologias de previsão de vazão com incorporação de precipitação, foram geradas e disponibilizadas pelo Centro de Previsão de Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE) as previsões meteorológicas para as bacias hidrográficas com aproveitamentos integrantes do SIN obtidas pelo modelo regional Eta (Cataldi *et al*., 2007), com resolução 40 km x 40 km, que apresenta resultados com qualidade razoavelmente satisfatória para os horizontes considerados (Guilhon *et al*.,2007).

Em 2006, foi autorizada pela ANEEL a utilização de um modelo complementar ao PREVIVAZ, o Modelo de Previsão de Classes de Vazão (MPCV), para a bacia do rio Iguaçu.

Como resultado dos estudos iniciados em 2005, a partir do PMO de junho de 2008, o ONS passou a utilizar outros modelos de previsão de vazões que incorporam também previsão de precipitação, com a devida anuência da ANEEL. Para a bacia do rio Paraná, no trecho incremental à UHE Itaipu, o modelo do tipo chuva-vazão *Soil Moisture Accounting Procedure* combinado com Modelagem Estocástica Linear (SMAP-MEL); para a bacia do alto/médio rio Grande até a UHE Porto Colômbia, o modelo SMAP; para a

bacia do rio Paranaíba, no trecho entre a UHE Itumbiara e a UHE São Simão, o Modelo de Grandes Barragens (MGB); para a bacia do rio Uruguai, o modelo MPCV; e, para a bacia do rio Iguaçu, o modelo FUZZY.

Entre 2010 e 2013, outros modelos entraram em operação para a previsão em trechos de diferentes bacias: o modelo NEURO3M, utilizado para previsão de vazões à UHE Três Marias (desde maio 2010); o SMAP para a bacia do rio Paranapanema (outubro de 2010); os modelos de previsão de vazões NeuroSF e CPINS para a previsão de vazões à UHE Sobradinho (março de 2012); o SMAP para as bacias do Alto Paranaíba até a UHE Itumbiara e do Baixo Grande entre Porto Colômbia e Água Vermelha, bem como foi autorizado o uso da metodologia de remoção de viés para a previsão da UHE Três Marias na bacia do rio São Francisco (setembro 2013).

A Tabela 3.1 resume algumas características dos modelos utilizados no setor elétrico. Alguns modelos serão abordados mais detalhadamente nos tópicos 3.3 e 3.4.

Tabela 3.1 - Modelos de previsão utilizados no setor elétrico até 2013. Fonte: elaboração própria, adaptado de www.ons.org.br.

Modelo	**Metodologia**	**Descrição**	**Aplicação (qtde. reservatórios)**	**Escala Temporal**	**Utiliza informação meteorológica**
MPCV	Redes Neurais e Treinamento Bayesiano	Item 3.4.1	Rio Uruguai (07)	Diária, até 12 dias a frente	Sim (modelo Eta regional)
SMAP-MEL	Modelagem conceitual + Modelagem estocástica	Item 3.4.2	Rio Paraná, no trecho incremental à UHE Itaipu (0)	Diária, até 12 dias a frente	Sim (modelo Eta regional)
SMAP	Modelagem conceitual	Item 3.4.3	Rio Grande (15); Alto Paranaíba até a UHE Itumbiara (08); Rio Paranapanema (10)	Diária, até 12 dias a frente	Sim (modelo Eta regional)
MGB [1]	Modelo distribuído de base conceitual	Item 3.4.4	Rio Paranaíba, no trecho entre a UHE Itumbiara e a UHE São Simão (02)	Diária, até 12 dias a frente	Sim (modelo Eta regional)
FUZZY	Fuzzy recorrente	Item 3.4.5	Rio Iguaçu (03)	Diária, até 12 dias a frente	Sim (modelo Eta regional)
NEURO3M [2]	Redes Neurais	-	Trecho afluente à UHE Três Marias (0)	Diária, até 12 dias a frente	Sim (modelo Eta regional)
CPINS [3]	Modelagem física (propagação)	Item 3.3.3	rio São Francisco (exceto UHE Três Marias) (05)	Diária, até 12 dias a frente	Não
PREVIVAZ [4]	Modelos estocásticos **univariados**	Item 3.3.1	**Restante das usinas do SIN** (87, de um total de 141)	Semanal (até 6 semanas a frente)	Não
PREVIVAZM		Item 3.3.1	Todas as usinas do SIN	Mensais (até 12 meses a frente)	Não
PREVIVAZH		Item 3.2.2	Apenas se necessário	Diárias	Não

Observações:

[1] Os Modelos MGB e SMAP são utilizados também na previsão da UHE Baguari, no rio Doce, e das UHEs Jupiá, Porto Primavera e Itaipu, no rio Paraná.

[2] Modelo utilizado para melhorar as previsões afluentes à UHE 3 Marias.

[3] Além do modelo CPINS, utiliza-se o modelo NeuroSF no trecho entre a UHE 3 Marias e a UHE Sobradinho.

[4] Independentemente da bacia hidrográfica, o modelo PREVIVAZ é utilizado para completar as previsões semanais de vazões, desde a segunda semana operativa até a previsão com seis semanas de antecedência.

3.3. MODELOS DE PREVISÃO DE VAZÃO DO SETOR ELÉTRICO QUE NÃO UTILIZAM INFORMAÇÕES METEOROLÓGICAS

3.3.1. Modelos PREVIVAZ E PREVIVAZM

O modelo PREVIVAZM (Costa *et al*., 2003) tem como objetivo fornecer previsões mensais de afluências a aproveitamentos hidroelétricos do Sistema Hidroelétrico Brasileiro para um horizonte de até 12 meses, sendo uma ferramenta para estudos de verificação de condições de atendimento da demanda energética no horizonte anual. Já o modelo PREVIVAZ foi desenvolvido visando à obtenção das previsões de afluências semanais, até seis semanas à frente, para serem utilizadas no primeiro mês do planejamento de curto prazo (PMO e suas revisões semanais). Este modelo é, portanto, executado ao final de cada mês para a elaboração do PMO do mês seguinte e, durante o mês em curso, é executado todas as semanas para a realização das previsões do PMO (Costa *et al*., 2007).

O PREVIVAZ e o PREVIVAZM utilizam modelos de séries temporais lineares do tipo PARMA (p,q) (Box e Jenkins, 1970), acoplados a diferentes pré-transformações das séries históricas, tipo Box-Cox (Box e Cox, 1964) e a diferentes formas de estimação de parâmetros dos modelos, para obtenção de previsões de vazões semanais e mensais respectivamente (Costa *et al*., 2007).

No desenvolvimento do PREVIVAZ, as formulações lineares de previsão de séries temporais foram divididas em duas classes de modelos: modelos estacionários e periódicos. Na classe de modelos estacionários estão a média anual, a média de cada semana e os modelos autoregressivos média-móvel ARMA(p,q), em que a equação de regressão e a estrutura de correlação de uma dada semana em relação às anteriores foram consideradas constantes ao longo do ano, tendo sido considerados na modelagem as médias e desvios-padrão sazonais (isto é, de cada semana). Para a estimação dos parâmetros desses modelos utiliza-se o método dos momentos (Costa *et al*., 2007).

Na classe de modelos periódicos utilizam-se os modelos PARMA(p,q), que se caracterizam por apresentar uma equação de regressão para cada período. Nos modelos PAR(p), a estimação dos parâmetros é feita pelo método dos momentos e também por regressão linear em relação à origem das previsões (regressão usando-se apenas as vazões semanais anteriores a origem da previsão). Já nos modelos PARMA(p,q), foram utilizados

o método dos momentos, regressão simples e regressão em relação a origem das previsões (Costa *et al*., 2007).

Uma alternativa também considerada pelos desenvolvedores do PREVIVAZ é a estrutura de correlação com sazonalidade mensal, trimestral e semestral. Esta facilidade, de acordo com Costa *et al.* (2007) "tem a vantagem de aumentar a quantidade de informação para a estimação da função de correlação". As diversas alternativas de classe de modelo, ordem dos modelos, agrupamento da estrutura de autocorrelação e métodos de estimação dos parâmetros resultaram em diferentes algoritmos de previsão, implementados no PREVIVAZ, conforme Tabela 3.2. Qualquer um dos algoritmos pode ser utilizado com aplicação da transformação Box-Cox, Logarítmica ou sem transformação da série histórica de afluência (Costa *et al*., 2007).

No modelo PREVIVAZ, os algoritmos de previsão, definidos por um modelo, método de estimação específico e transformação da série de vazões, são testados por um esquema de validação cruzada, onde cada série é dividida em duas partes. A primeira parte da série é utilizada para a estimação dos parâmetros (estimação) e a segunda parte, para o cálculo de erros de previsão (verificação). Em seguida, a estimação dos parâmetros passa a ser feita com a segunda parte da série, ficando a primeira parte para cálculo de erros de previsão. Para cada parte da série é computado o erro padrão de previsão um passo à frente (por meio do indicador raiz quadrada da média dos quadrados dos erros de previsão) obtendo-se, a seguir, a média dos dois valores. O modelo PREVIVAZ utiliza a cada semana um dos algoritmos enumerados na Tabela 3.2, sendo escolhido aquele que gerou as previsões um passo a frente com o menor erro médio quadrático de previsão (Costa *et al*., 2007).

O modelo PREVIVAZM foi desenvolvido a partir do modelo PREVIVAZ e segue a mesma abordagem, adaptada ao intervalo mensal e ao horizonte de 12 meses, mantendo-se o procedimento de estimação e escolha do melhor algoritmo de previsão. A Tabela 3.3 apresenta os algoritmos de previsão implementados no modelo PREVIVAZM. Estes algoritmos podem também ser aplicados às séries históricas de vazões mensais com transformação Box-Cox e nas séries sem transformação (Costa *et al*., 2007).

Tabela 3.2 - Modelos implementados no PREVIVAZ. Fonte: Costa *et al.* (2007).

Modelo de Previsão	**Característica**	**Método de Estimação**
CONSTANTE	Previsão pela média anual	momentos
SAZONAL	Previsão pela média da semana	momentos
AR (p) ($1 \leq P \leq 4$)	Estrutura de correlação estacionária	momentos
ARMA (p,1) ($1 \leq P \leq 3$)	Estrutura de correlação estacionária	momentos
PAR (p) – G(1) ($1 \leq P \leq 4$)	Estrutura de correlação semestral	momentos
PAR (p) – G(2) ($1 \leq P \leq 4$)	Estrutura de correlação trimestral	momentos
PAR (p) – G(3) ($1 \leq P \leq 4$)	Estrutura de correlação mensal	momentos
PAR (p) – G(4) ($1 \leq P \leq 4$)	Estrutura de correlação semanal	momentos
PAR (p) – R0 ($1 \leq P \leq 4$)	Estrutura de correlação sazonal	Regressão-origem
PARMA (p,1) – G(1) ($1 \leq P \leq 4$)	Estrutura de correlação semestral	momentos
PARMA (p,1) – G(2) ($1 \leq P \leq 4$)	Estrutura de correlação trimestral	momentos
PARMA (p,1) – G(3) ($1 \leq P \leq 4$)	Estrutura de correlação mensal	momentos
PARMA (p,1) – G(4) ($1 \leq P \leq 4$)	Estrutura de correlação semanal	momentos
PARMA (p,1) – R0 ($1 \leq P \leq 3$)	Estrutura de correlação sazonal	Regressão-origem
PARMA (p,1) – R ($1 \leq P \leq 3$)	Estrutura de correlação sazonal	Regressão

Tabela 3.3 - Modelos implementados no PREVIVAZM. Fonte: Costa *et al.* (2007).

Modelo de Previsão	**Característica**	**Método de Estimação**
CONSTANTE	Previsão pela média anual	momentos
SAZONAL	Previsão pela média do mês	momentos
AR (p) ($1 \leq P \leq 4$)	Estrutura de correlação estacionária	momentos
ARMA (p,1) ($1 \leq P \leq 3$)	Estrutura de correlação estacionária	momentos
PAR (p) – G(1) ($1 \leq P \leq 4$)	Estrutura de correlação semestral	momentos
PAR (p) – G(2) ($1 \leq P \leq 4$)	Estrutura de correlação trimestral	momentos
PAR (p) – G(3) ($1 \leq P \leq 4$)	Estrutura de correlação mensal	momentos
PARMA (p,1) – G(1) ($1 \leq P \leq 4$)	Estrutura de correlação semestral	momentos
PARMA (p,1) – G(2) ($1 \leq P \leq 4$)	Estrutura de correlação trimestral	momentos
PARMA (p,1) – G(3) ($1 \leq P \leq 4$)	Estrutura de correlação mensal	momentos
PARMA (p,1) – R ($1 \leq P \leq 3$)	Estrutura de correlação sazonal	Regressão simples

Os modelos PREVIVAZ e PREVIVAZM não possuem em sua estrutura uma relação de dependência entre si, ou seja, o resultado das previsões de um modelo não exerce influência nas previsões geradas pelo outro modelo. A não existência de uma estrutura hierárquica temporal de previsões pode resultar na redução da qualidade das previsões, uma desvantagem da utilização dos modelos da família PREVIVAZ.

3.3.2. Modelo PREVIVAZH

O PREVIVAZH (Livino de Carvalho, 2001) é um modelo de natureza estocástica, baseado na desagregação em intervalos diários das previsões semanais utilizadas na elaboração do PMO (em geral, obtidas pelo modelo PREVIVAZ). O modelo utiliza as previsões das afluências semanais e sequências sintéticas de vazões diárias geradas pelo modelo DIANA (Kelman *et al*., 1983) para os sete dias da semana, condicionada às últimas afluências diárias observadas. A geração de seqüências sintéticas de afluências diárias se baseia no modelo DIANA, alterado para condicionar a geração de seqüências às informações hidrológicas recentes, ou seja, às duas últimas afluências diárias observadas. O modelo DIANA considera que as vazões diárias podem ser decompostas em duas parcelas: uma parcela dependente de fatores externos (por exemplo, precipitação) e outra que representa o contínuo esvaziamento da bacia. (Costa *et al*., 2007).

3.3.3. Modelo CPINS

O modelo de previsão CPINS é utilizado na operação diária em tempo real e para um horizonte de programação de até 14 dias para a bacia do rio São Francisco (Paiva e Acioli, 2007).

O modelo CPINS utiliza a rotina de propagação do modelo SSARR – *Streamflow Synthesis and Reservoir Regulation (Corps of Engineers*, 1967) para a propagação das vazões nos diversos trechos de rio e entre os reservatórios, a qual estabelece que o amortecimento da hidrógrafa de montante em relação à de jusante originado ao longo do rio é calculado através da equação da continuidade, subdividindo o trecho de rio num certo

número de reservatórios fictícios (Paiva e Acioli, 2007). Após algumas transformações algébricas, tem-se:

$$O_2 = O_1 + \left(\frac{I_m - O_1}{T_s - t/2}\right) t \tag{3.1}$$

A Equação (3.1) é a expressão básica da rotina de propagação do SSARR, onde:

I_m = média das afluências

O_1 = defluência no início do período

O_2 = defluência no fim do período

t = intervalo de tempo de alimentação

T_s = tempo de armazenamento

O tempo de armazenamento para propagação em canais pode ser especificado como uma tabela "T_s" versus descargas de jusante, ou ser informado como sendo uma função das descargas. Normalmente, quando o fluxo está confinado em canal, o tempo de armazenamento obedece à seguinte relação:

$$T_s = \frac{C_k}{Q^{Cn}} \tag{3.2}$$

em que Q é a média entre as defluências de início e de fim de período

$$Q = \frac{O_2 + O_1}{2} \tag{3.3}$$

T_s = tempo de armazenamento

C_k = constante calibrada a partir de dados observados

C_n = coeficiente calibrado a partir de dados observados e que normalmente varia entre 1 e –1.

3.4. MODELOS DE PREVISÃO DE VAZÃO QUE UTILIZAM INFORMAÇÕES CLIMÁTICAS

Conforme estudos de Guilhon *et al.* (2007), a análise de desempenho do modelo PREVIVAZ no período de 1997 a 2004, considerando uma semana a frente, mostrou que

as vazões médias semanais previstas apresentaram desvios absolutos em relação às vazões naturais observadas, de 12% a 68%. No caso da previsão de vazões médias diárias, considerando um dia à frente, os desvios do modelo PREVIVAZH ficaram entre 10% e 52%, para os anos 2000 e 2001. A Tabela 3.4 apresenta os desvios médios absolutos percentuais dos modelos de previsão de vazões naturais médias semanais (PREVIVAZ) e diárias (PREVIVAZH) em alguns aproveitamentos do SIN.

Tabela 3.4 - Desvios médios absolutos de previsão de vazões dos modelos PREVIVAZ (entre 1997 e 2004) e PREVIVAZH (entre 2000 e 2001). Fonte: Guilhon *et al.* (2007).

Usina	**Desvio Previvaz (%)**	**Desvio Previvazh (%)**
Tucuruí	12,3	10,2
Itaipu	15,5	30,4
São Simão	18,4	17,7
Marimbondo	20,9	16,2
Salto Grande	21,6	12,9
Furnas	27,9	24,2
Paraibuna	28,6	37,2
Três Marias	35,3	31,8
Serra da Mesa	35,4	51,7
Foz do Areia	38,0	12,2
Passo Real	42,0	27,9
Itá	67,6	Não avaliado

Os elevados desvios apresentados pelos modelos motivaram a criação de um grupo de trabalho no âmbito do ONS com o objetivo de analisar e propor aperfeiçoamentos dos modelos computacionais implantados no planejamento e programação da operação, bem como estabelecer metodologias complementares para apoio à decisão operativa.

No decorrer do ano de 2005, o ONS contratou o desenvolvimento de diferentes tecnologias para a previsão de vazões diárias, para um horizonte de 12 dias à frente, incorporando, além das vazões naturais passadas, as informações da precipitação observada e prevista.

Considerando o melhor desempenho dos modelos que utilizam informação meteorológica, quando comparado com os resultados da aplicação da metodologia baseada no sistema PREVIVAZ, foi autorizada pelo órgão regulador (ANEEL) a utilização dos novos modelos pelo ONS, em seus processos de planejamento e programação da operação (ONS, 2007a).

Em virtude das complexidades de se estender as previsões de chuva para horizontes de tempo superiores a uma semana, os modelos de previsão com uso de informação meteorológica ficam limitados temporalmente. A complementação das previsões com horizontes de duas a seis semanas a frente ainda é realizada com o modelo PREVIVAZ, que não incorpora nenhuma informação meteorológica. A seguir, serão apresentados alguns modelos utilizados pelo ONS que utilizam informações de precipitação.

3.4.1. Modelo de Previsão de Classes de Vazão (MPCV)

O modelo MPCV foi desenvolvido com metodologia baseada em uma previsão de faixas de vazão, como balizamento para a escolha do modelo do sistema PREVIVAZ, ou seja, utiliza classificações associadas à previsão direta de vazões naturais e verifica qual a previsão, dentre as realizadas pelo PREVIVAZ, estaria dentro da faixa prevista, sendo esta, então, considerada como a melhor previsão do sistema PREVIVAZ para determinado dia de previsão (Guilhon, 2003).

Sabendo-se que a vazão da semana em curso está dentro de determinada faixa de valores e a chuva prevista para a semana seguinte se encontrará numa outra determinada faixa de valores, é possível então prever, a partir de uma análise das probabilidades condicionais do histórico de dados, qual será a faixa mais provável em que se encontrará a vazão da semana seguinte. Para obter tal classificação, foram empregados algoritmos de Redes Neurais baseados em treinamento bayesiano (ONS, 2007a).

3.4.2. Modelo SMAP-MEL

A metodologia proposta pela Fundação Centro Tecnológico de Hidráulica (FCTH/USP) foi a escolhida para ser implementada no trecho incremental à Itaipu, entre as

UHEs Porto Primavera e Rosana, e inclui a combinação de duas modelagens distintas (ONS, 2007b):

- Modelagem hidrológica conceitual, com o emprego do modelo diário do tipo chuva-vazão conhecido como SMAP – *Soil Moisture Accounting Procedure* (Lopes *et al*., 1982);
- Modelagem estocástica linear (MEL), modelo estocástico linear multivariado, também considerando informações de precipitação, desenvolvido pela FCTH/USP.

Para a modelagem do SMAP diário, a bacia foi dividida em 7 sub-bacias, 4 das quais possuíam postos fluviométricos de controle, além da sub-bacia de jusante, afluente ao lago de Itaipu. O modelo hidrológico foi calibrado para as bacias que possuíam dados fluviométricos, transferindo-se os parâmetros para as bacias não controladas.

Para a modelagem MEL, a bacia foi dividida em quatro sub-bacias. O modelo foi calibrado para a bacia incremental entre as três bacias de montante, correspondentes a três postos fluviométricos e o reservatório de Itaipu, fazendo o melhor uso possível dos dados observados a montante, enquanto a sua antecedência permitia (menores horizontes de previsão). Para maiores horizontes, foram também desenvolvidos três modelos estocásticos para os postos fluviométricos de montante, de modo a injetar essas previsões, devidamente propagadas, no modelo de jusante.

Ao final, como os resultados dos dois modelos (SMAP e MEL) foram estatisticamente equivalentes, optou-se por uma solução mista, onde os resultados obtidos dos dois modelos são somados por meio de uma combinação linear. Os coeficientes (pesos) da combinação linear são calculados, a cada semana, em função da eficiência de cada modelo em fornecer isoladamente a melhor previsão nos últimos anos (ONS, 2007b).

3.4.3. Modelo SMAP

No modelo utilizado pelo ONS, foram realizadas algumas alterações em relação à formulação tradicional do modelo SMAP diário. Uma destas alterações refere-se ao ajuste na configuração do modelo para operação com quatro reservatórios, conforme pode ser visto na Figura 3.1. Justifica-se a opção por esse ajuste, pois, para vazões relativamente baixas, não há extravasamentos pelas margens e escoamento pelas planícies.

Outra vantagem da utilização do esquema de quatro reservatórios é que, caso queira-se inibir o funcionamento do quarto reservatório, basta fixar um valor relativamente elevado para o parâmetro H, referente à altura representativa para início de escoamento em planícies (mm). Dessa forma, pode-se utilizar o esquema do modelo SMAP diário apresentado na Figura 3.1 para qualquer sub-bacia.

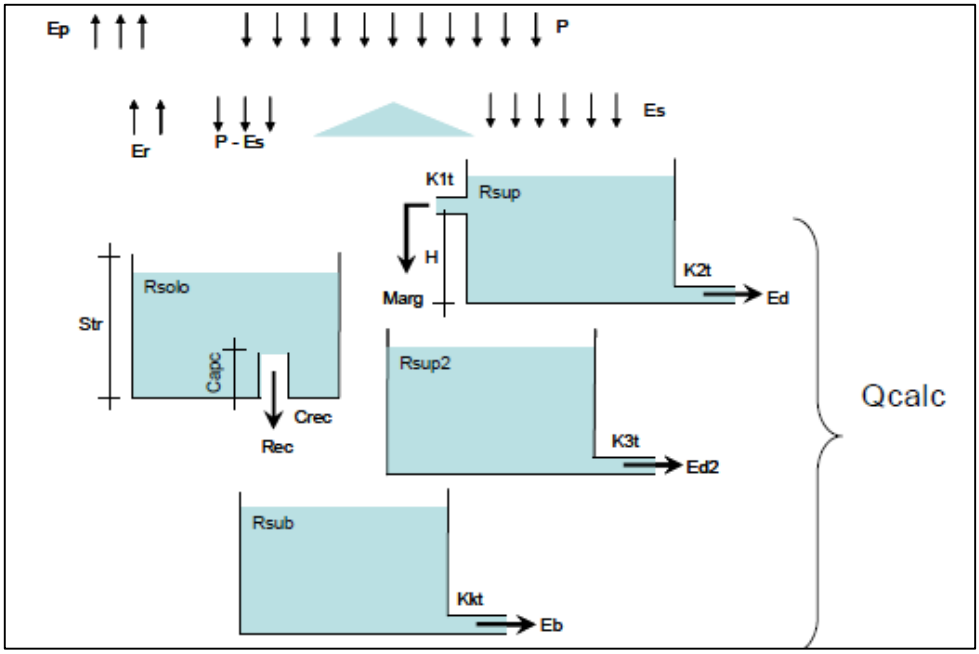

Figura 3.1 - Esquema do Modelo SMAP aplicado para a bacia do Rio Grande com 4 reservatórios. Fonte: ONS (2008)

3.4.4. Modelo MGB

Para a previsão de vazões na bacia incremental a São Simão, está sendo utilizado o Modelo de Grandes Bacias (MGB), um modelo hidrológico conceitual, distribuído, desenvolvido por Collischonn e Tucci (2001). A topologia da bacia é representada pelo modelo através de uma estrutura que determina, em função da topografia da bacia, o caminho que a água percorre em uma determinada célula.

De acordo com Jímenez e Collischonn (2015), o modelo, na sua versão mais recente, "discretiza a bacia em unidades irregulares, denominadas minibacias, que são definidas a partir de dados do relevo de um Modelo Digital de Elevação. O modelo é composto por módulos que estimam diferentes processos hidrológicos relevantes na transformação da chuva em vazão, como a interceptação, o balanço hídrico do solo, a

geração do escoamento em cada minibacia, a evapotranspiração e a propagação da vazão através da rede de drenagem."

Conforme Meller et al. (2014), o modelo MGB estima as vazões por meio de "procedimento de assimilação de dados apresentado nos trabalhos de Paz et al. (2007) e Collischonn et al. (2007). Esse procedimento utiliza as vazões observadas em tempo-real para atualizar as condições iniciais do modelo, representadas pelas vazões calculadas ao longo da rede de drenagem e o volume de água armazenado nos reservatórios superficial, subsuperficial e subterrâneo de cada minibacia do modelo hidrológico."

3.4.5. Modelo Fuzzy

A metodologia baseia-se na técnica de Extração de Conhecimento em Bases de Dados (ou *Knowledge Discovery in Databases*), que é o processo de extrair informação válida e previamente desconhecida, a partir de grandes bases de dados. O processo de extração de conhecimento envolve diversas etapas como: definição do problema, pré-processamento de dados, mineração e interpretação de dados (Evsukoff *et al*., 2007).

O modelo Fuzzy utilizado no rio Iguaçu trata-se, na verdade, de dois modelos: um modelo de dados e um modelo de previsão. O modelo de dados, denominado cubo OLAP (On-Line Analytical Procedure), permite a visualização da informação armazenada sob diversas perspectivas e níveis de agregação. Esta tecnologia exige que os dados não contenham falhas nos registros, fato que demanda o desenvolvimento de um modelo de preenchimento de falhas nos dados básicos. O modelo de previsão é denominado modelo fuzzy recorrente, cujos detalhes podem ser obtidos em Evsukoff e Ebecken (2004).

3.5. VARIABILIDADE CLIMÁTICA E PADRÕES DE CHUVA E VAZÃO NO BRASIL

O gerenciamento de recursos hídricos envolve o entendimento da variabilidade do sistema climático. Previsões de chuvas e de vazões confiáveis em longo prazo podem melhorar a gestão de sistemas de recursos hídricos (Araghinejad *et al*., 2006).

A variabilidade temporal dos fenômenos climáticos condiciona o regime de precipitação e evaporação em uma bacia hidrográfica, e desta forma, determina os padrões

de variação das vazões. Compreender a correlação entre a precipitação e as anomalias de vazão, e as alterações nos padrões oceano-atmosféricos de larga escala resultam em melhorias na previsão de longo prazo (Hamlet e Lettenmaier, 1999; Chiew *et al*, 1998;. Sharma, 2000a; apud Araghinejad *et al*., 2006).

3.5.1. Padrões de teleconexão

A palavra teleconexão significa conexão a distância e, em meteorologia, explica como anomalias que ocorrem em uma região são associadas a anomalias em regiões remotas. O termo "padrão de conexão", ou simplesmente "teleconexão", de acordo com Cavalcanti *et al*. (2009) "refere-se a um padrão recorrente e persistente de anomailas de uma determinada variável, por exemplo, pressão e circulação de grande escala, que cobre vastas áreas geográficas. Padrões de teleconexão são também conhecidos como modelos preferenciais de variabilidade de baixa frequência. Embora esses padrões possam persistir por várias semanas ou meses, algumas vezes eles podem se tornar dominantes por vários anos consecutivos e, dessa forma, mostram uma parte importante da variabilidade interanual e interdecadal da circulação atmosférica".

Alguns padrões podem ser representados através de índices climáticos que quantificam a sua intensidade, como por exemplo, os índices NINO 1+2, NINO 3, NINO 4 e NINO 3.4, que representam a Temperatura da Superfície do Mar (SST) em quatro regiões do Oceano Pacífico, o índice PDO (*Pacific Decadal Oscillation Index*) que representa a anomalia da SST no Oceano Pacífico, o índice SOI (Southern Oscillation Index), o índice AMO (*Atlantic Multi-decadal Oscillation*), dentre outros. No item 3.5.2 são apresentados os fenômenos El Niño/La Niña e os índices climáticos NINO 1+2, NINO 3, NINO 4, NINO 3.4.

3.5.2. El Niño/La Niña - Oscilação Sul

El Niño é um termo geral usado para descrever um conjunto de eventos climáticos incomuns e concorrentes no sul do Pacífico que estão correlacionados com eventos extremos hídricos em diferentes partes do mundo. O termo El Niño originalmente referia-se à água de superfície relativamente quente que aparece ao longo da costa oeste equatorial da América do Sul durante os primeiros meses do ano, devido a um enfraquecimento anual

dos ventos alísios. Agora significa um aquecimento generalizado, em comparação com a média, nas regiões central e oriental do Oceano Pacífico equatorial. Ao mesmo tempo, as temperaturas da superfície do mar no Pacífico ocidental são mais frias do que a média (Kelman *et al*., 2000).

Em anos normais, a SST da parte ocidental do Oceano Pacífico equatorial fica mais quente do que a parte oriental e a pressão na parte oriental do Oceano Pacífico é maior do que a da parte ocidental. Durante anos anômalos, a SST da parte oriental do Oceano Pacífico equatorial torna-se mais quente do que a normal e o campo de pressão é invertido, ou seja, a pressão anormal na parte oriental do Oceano Pacífico torna-se menor do que na parte ocidental, podendo também ocorrer a situação contrária (Maity e Kumar, 2008).

O aquecimento (resfriamento) anômalo da SST sobre a parte leste do Oceano Pacífico é conhecido como El Niño (La Niña), enquanto a variação anômala da interação oceano-continente do campo de pressão, entre as partes oriental e ocidental do Oceano Pacífico, é chamada de Oscilação Sul. Conforme Maity e Kumar (2008), atuando em conjunto, as partes oceânicas e atmosféricas são conhecidas em conjunto como o fenômeno El Niño Oscilação Sul (ENOS).

O fenômeno ENOS é uma perturbação interanual do sistema climático, caracterizada pelo enfraquecimento dos ventos alísios e aquecimento das águas superficiais do Oceano Pacífico Equatorial. Atualmente o El Niño é definido como um aumento de pelo menos 0,5^{o}C da SST na região leste-central equatorial do Oceano Pacífico Sul durante um período mínimo de 3 meses, ocorrendo em intervalos irregulares de 2 a 7 anos. A La Niña é o processo inverso, ou seja, é a redução da SST durante 3 meses em pelo menos 0,5^{o}C, que ocorre em intervalos de 3 a 5 anos (Climate Prediction Center, 2013).

O ENOS indica o aquecimento anômalo de larga escala da SST sobre o Oceano Pacífico central e leste, com mudança associada ao campo de pressão (Maity e Kumar, 2008).

Em anos normais, observa-se a formação da célula de Walker de circulação atmosférica, caracterizada por movimentos ascendentes no Pacífico Central/Oriental e movimentos descendentes na região oriental da América do Sul, com ventos na alta atmosfera indo do oeste para o leste e com ventos superficiais (alísios) indo do leste para o oeste (Figura 3.2).

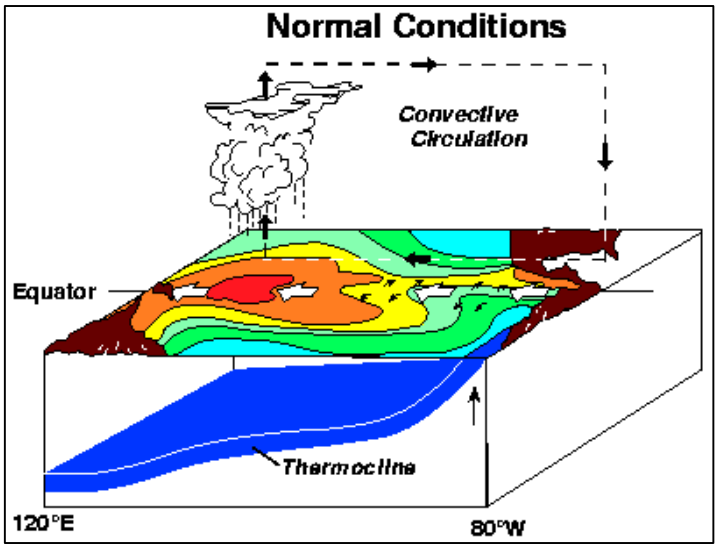

Figura 3.2 - Oceano Pacífico em condições normais

Fonte: http://www.pmel.noaa.gov/tao/elnino/nino_normal.html#normal (acesso em 05/12/2014).

Nos anos de El Niño ocorre a bifurcação da célula de Walker, com os ventos alísios tornando-se menos intensos, associada com o aumento da temperatura na superfície do Pacífico Tropical (Figura 3.3).

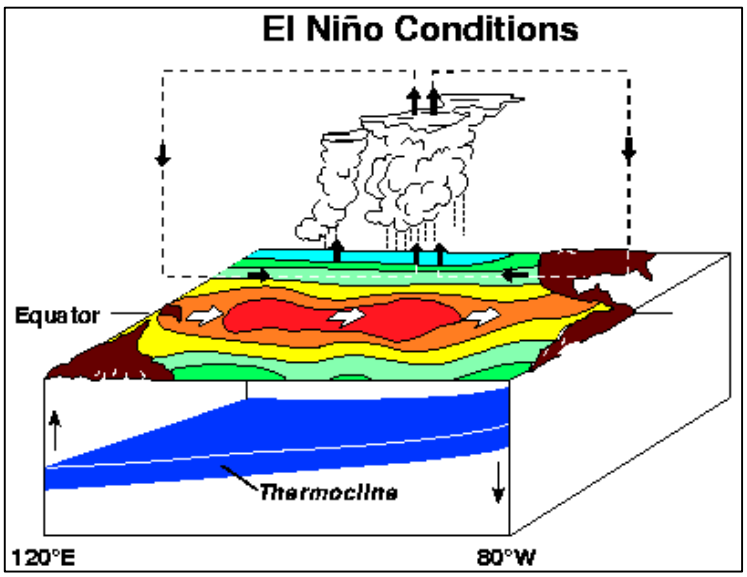

Figura 3.3 - Oceano Pacífico em condições de El Niño

Fonte: http://www.pmel.noaa.gov/tao/elnino/nino_normal.html#nino (acesso em 05/12/2014).

Nos anos de La Niña os ventos alísios tornam-se mais intensos e a SST do Oceano Pacífico se torna mais baixa que a habitual (Figura 3.4).

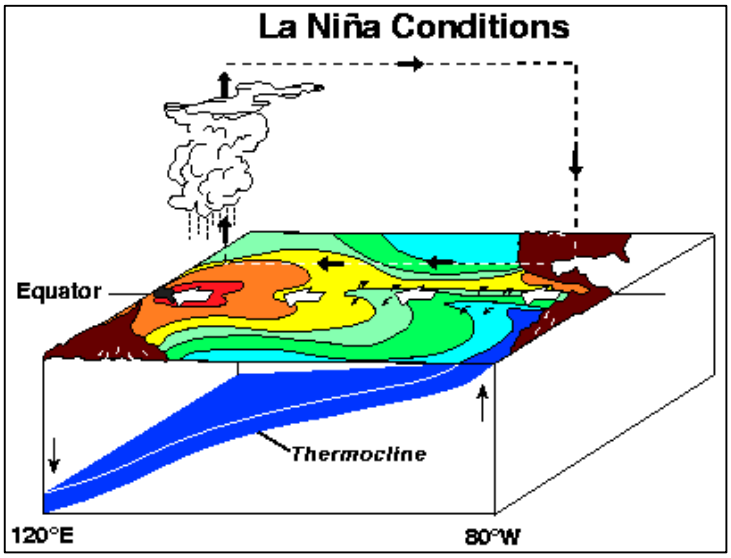

Figura 3.4 - Oceano Pacífico em condições de La Niña

Fonte: http://www.pmel.noaa.gov/tao/elnino/nino_normal.html#nina (acesso em 05/12/2014).

De forma a monitorar as alterações na SST do Oceano Pacífico foram criados quatro índices: NINO 1+2, NINO 3, NINO 4 e NINO 3.4. Estes índices correspondem às SST nas regiões da costa oriental da América do Sul (0 a 10ºS e 90 a 80ºW) (NINO 1+2), Pacífico Ocidental/Central (5ºS a 5ºN e 150 a 90ºW) (NINO 3), Pacífico Oriental/Central (5ºS a 5ºN e 160ºE a 150ºW) (NINO 4) e Pacífico Central (5ºS a 5ºN e 170 a 120ºW) (NINO 3.4), conforme Figura 3.5.

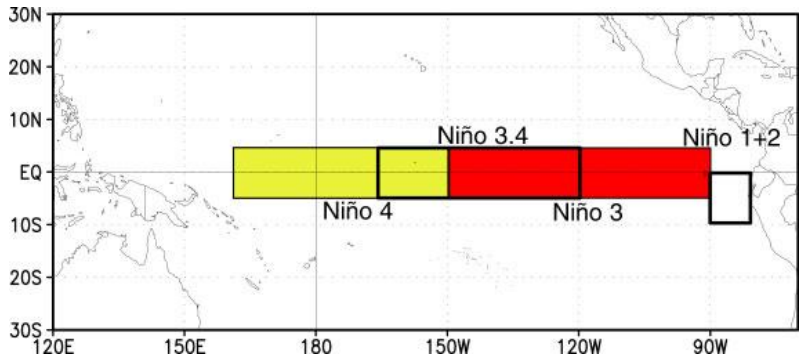

Figura 3.5 - Regiões do ENOS

Fonte: http://oceanworld.tamu.edu/resources/oceanography-book/equatorialcurrents.htm. (acesso em 05/12/2014).

3.5.3. Mudanças nos padrões de chuva e vazão associadas com eventos de El Niño no Brasil.

Entre os impactos mais conhecidos de El Niño na costa oeste da América do Sul, destaca-se o aumento da precipitação sobre o Equador e o Peru e a redução de precipitação sobre a maior parte da Colômbia (Kelman *et al*., 2000).

No Brasil, os eventos de El Niño geralmente levam a um padrão dipolo de chuvas, condições mais secas no Nordeste do Brasil durante a primavera austral e no verão e condições mais úmidas, consistentes principalmente na região sul do Brasil, na primavera austral (Ropelewski e Halpert, 1987; Grimm *et al*, 1998; apud Lima e Lall, 2010).

Particularmente, o aumento da precipitação sobre a região sul tem sido associado a jatos subtropicais de oeste, 200-hPa acima do normal durante eventos de El Niño, que por sua vez, favorecem a intensificação dos centros de convecção de mesoescala no sul do Brasil. Ele também desloca uma circulação anti-ciclônica anômala pro sudeste do Brasil e uma anomalia ciclônica pro sudoeste da América do Sul que favorecem instabilidades baroclínicas e padrões anômalos de precipitação no sul do Brasil (Ropelewski e Halpert, 1987; Cardoso e Dias, 2006; Coelho *et al*, 2002; Diaz e Markgraf, 2000; Grimm, 2004; Grimm *et al*, 2000, 1998; apud Lima e Lall, 2010). Grimm (2003, 2004) também concluiu que os efeitos do ENOS na precipitação sobre a região centro-oeste tendem a ser suavizados em uma base analítica sazonal e mostram que a região experimenta uma anomalia positiva (negativa) de chuva em maior parte no Janeiro subsequente a um evento de El Niño (La Nina). A Figura 3.6 ilustra os padrões típicos de alteração no comportamento da chuva e temperatura em diferentes partes do globo, entre dezembro e fevereiro, durante a ocorrência de El Niño.

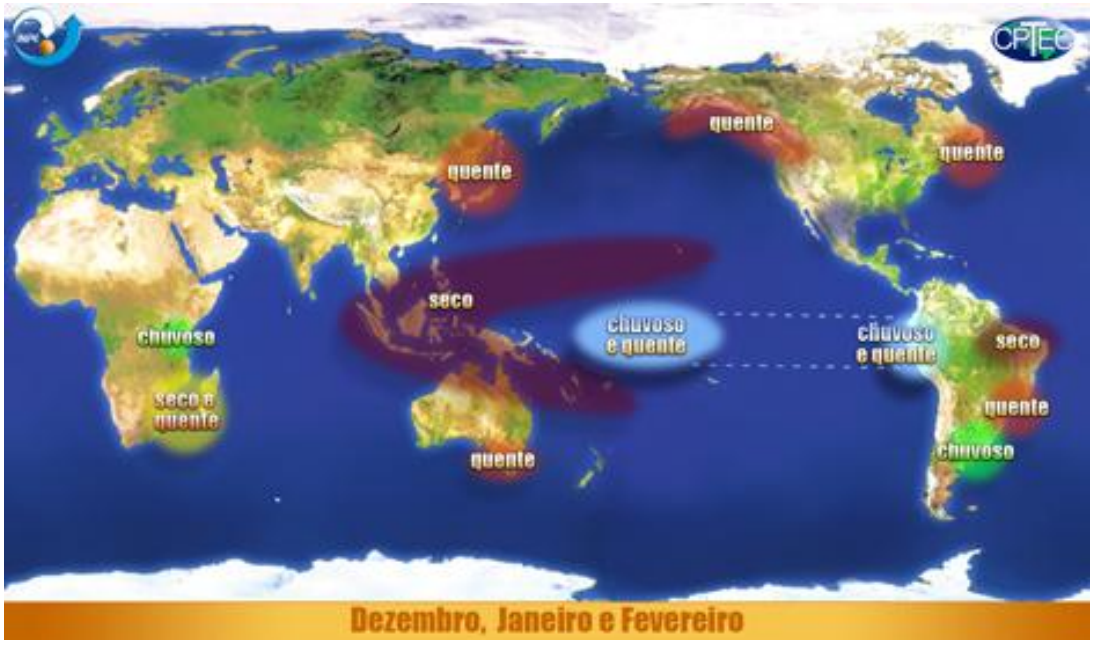

Figura 3.6 - Típicos padrões de chuva e temperatura associados com a ocorrência de El Niño. (Fonte: CPTEC, 2014).

3.5.4. A zona de convergência do Atlântico Sul (ZCAS).

As chuvas de verão austral nas partes central, sudeste e sul do Brasil estão associadas com a ZCAS, que resulta de um padrão de circulação dos ventos e umidade semelhante à circulação de sistemas de monção. A alta do Atlântico sul subtropical e o centro de baixa pressão continental (sul do Brasil e da Argentina) induzem a um giro na escala continental responsável pelo transporte de umidade do Oceano Atlântico para a região amazônica e posteriormente para o sul até os subtrópicos e latitudes médias da América do Sul (Carvalho *et al*, 2004;. Vera *et al*, 2006; apud Lima e Lall, 2010).

A ZCAS é caracterizada por uma banda orientada noroeste-sudeste de precipitação máxima que se estende ao longo da costa sudeste no Atlântico Sul (Lenters e Cook, 1995). A precipitação ao longo desta zona é impulsionada por pulsos de calor latente da bacia Amazônica, por efeitos orográficos da cordilheira dos Andes e por anomalias na SST (Lenters e Cook, 1995; Figueroa *et al.*, 1995; Berbery e Collini, 2000). Melhorias na ZCAS resultaram em inundações no sudeste e parte do sul do Brasil (Lima e Lall, 2008-b) e também têm sido associadas com eventos de El Niño (Liebmann *et al.*, 1999;. Barros *et al.*, 2000;. Carvalho *et al*., 2004; apud Lima e Lall, 2010).

O principal sistema gerador de precipitações sobre a região equatorial do Oceano Atlântico é a Zona de Convergência Intertropical (ZCIT), como foi mostrado por Hastenrath e Heller (1977) e Citeau *et al.* (1988).

3.6. USO DE INFORMAÇÃO METEOROLÓGICA EM PREVISÕES DE VAZÃO

A variabilidade climática oceano-atmosférica e a SST podem fornecer informações preditivas importantes sobre a variabilidade hidrológica nas regiões ao redor do mundo (Tootle *et al*., 2008). O entendimento físico de como fenômenos climáticos remotos (de larga escala) influenciam na variabilidade das vazões é fundamental para a construção dos modelos estatísticos que irão incorporar as informações meteorológicas.

Abordagens estatísticas que relacionam a vazão às variáveis oceano-atmosféricas de larga escala fornecem uma base para previsões eficientes – desde escalas sazonais a interanuais (Sharma, 2000a).

A relação entre chuva e vazão e os sinais oceano-atmosféricos tem sido observada em muitas partes do mundo. Conforme citam Araghinejad *et al.* (2006), exemplos são encontrados nas pesquisas de Chiew *et al.* (1998) para chuvas e vazões na Austrália, de Gutiérrez e Dracup (2001) para vazões na Colômbia, de Fowler e Kilsby (2002) para a região norte da Inglaterra, de Harshburger *et al.* (2002) para chuvas e vazão em Idaho, e de Karamouz e Zahraie (2004) para a Bacia do Rio Salt, no Arizona.

O relacionamento entre os sinais do tempo e de vazão é geralmente não-linear devido à complexa dinâmica da interação oceano-atmosfera com climas regionais. Além disso, há incertezas em determinar os preditores de vazão em modelos de previsão de longo prazo. Devido à complexidade e à não-linearidade do processo, é desejável a aplicação de um método probabilístico de previsão (Araghinejad *et al*., 2006).

Algumas anomalias climáticas sazonais são associadas com a variação da SST em escalas que vão de sazonais a interanuais (Tippett *et al*., 2004). Um exemplo da conexão entre a precipitação sazonal e a SST são anomalias de precipitação associadas com ENOS (Ropelewski e Halpert, 1987; Mason e Goddard, 2001).

Muitos pesquisadores têm investigado a relação estatística entre as variáveis hidroclimáticas e os sinais oceano-atmosféricos, tais como ENOS e a Oscilação do Atlântico Norte (NAO), para uso em modelos de previsão de longo prazo.

Tootle at al. (2008) identificaram que os fenômenos ENOS, PDO e AMO influenciam na hidrologia, citando como exemplos os trabalhos de Cayan e Peterson, 1989; Cayan e Webb, 1992; Kahya e Dracup, 1993, 1994a, b; Enfield *et al*, 2001; Rogers e Coleman, 2003; Maurer *et al*., 2004, McCabe *et al*., 2004.

O valor de precipitação sazonal é visto como uma variável aleatória com distribuição climatológica. Esta distribuição climatológica é então comparada com a distribuição referente ao valor de precipitação para uma particular SST. O impacto da SST na precipitação sazonal é medido pelo grau em que as duas distribuições diferem. Se a SST não tem impacto na precipitação, as duas distribuições serão idênticas. Caso contrário, as duas distribuições diferem significativamente (Tippett *et al*., 2004).

Embora cada um desses fenômenos oceânico-atmosféricos represente a variabilidade da SST, essa variabilidade é representada por uma região específica, espacialmente pré-determinada. No entanto, a utilização de SST para regiões inteiras (Pacífico e Atlântico) eliminaria qualquer viés espacial que a região (ou regiões) oceânica(s) de SST impacta(m) na hidrologia, auxiliando na identificação de regiões que não podem ser representadas por índices existentes e que poderiam conduzir a melhores previsões (Tootle at al., 2008).

Por natureza, os dados geofísicos como vazões ou temperaturas são muitas vezes sazonalmente estacionários. Ou seja, são estatisticamente semelhantes de ano para ano, mas podem variar drasticamente ao longo das estações. A sazonalidade dos dados acrescenta um grau de complexidade para a seleção e desenvolvimento de um modelo estocástico (Noakes *et al*., 1985).

Tippett *et al.* (2004) identificaram que variações nas escalas temporais (sazonais a interanuais) na SST conduzem a alterações nas estatísticas meteorológicas sazonais e anomalias climáticas sazonais.

3.6.1. Alguns Resultados

Hamlet e Lettenmaier (1999) incorporaram os preditores ENOS e PDO em uma abordagem conjunta de previsão de vazões. Eles definiram categorias climáticas para diferentes fases de ENOS e PDO para prever regimes climáticos nessas categorias. Eles forneceram diferentes previsões de vazões de até 6 meses para o rio Columbia e usaram

um modelo hidrológico de macroescala na bacia do rio Columbia para gerar hidrogramas de vazão.

Kelman *et al.* (2000) adotaram um modelo auto-regressivo utilizando a informação El Niño para refinar o processo de estimação de parâmetros para cada intervalo de tempo. A metodologia de combinação linear de valores de vazão foi testada para as séries temporais mensais de energias afluentes às hidroelétricas dos dezoito rios mais importantes da Colômbia.

Sharma (2000b) utilizou técnicas não paramétricas de estimação de densidade *kernel* para estimar a distribuição condicional de probabilidade de chuva. A abordagem baseou-se no uso de métodos *kernel* não paramétricos para estimativa de densidade de probabilidade uni e multivariada. Sharma (2000a) aplicou essa abordagem na identificação preditiva das chuvas trimestrais na barragem de Warragambana, na Austrália por meio de indicadores ENOS. Sharma (2000b) propôs um preditor único para as previsões ao invés da combinação de preditores.

Anderson *et al.* (2001) utilizaram um modelo numérico capaz de representar as principais características do sistema hidroclimático. A resposta hidrológica para os eventos ENOS foi simulada pelo método Monte Carlo, aplicando as condições iniciais aleatórias para obter uma ampla gama de possíveis respostas ao sistema hidrológico resultantes de um evento ENOS.

Piechota *et al.* (2001) desenvolveram um modelo de previsão de vazões com probabilidade de excedência utilizando vários preditores e aplicaram-na a cinco bacias hidrográficas australianas. Eles usaram como preditores o Índice de Oscilação Sul (SOI), a SST e a vazão do ciclo hidrológico do período anterior (persistência). A metodologia de previsão, denominada Análise do Discriminante Linear (LDA), produz a probabilidade contínua de excedência para cada preditor e, em seguida, determina uma combinação linear através de um processo de otimização, que determina o peso de cada modelo para se produzir a máxima qualidade nas previsões finais.

Souza Filho e Lall (2003) realizaram uma investigação da dependência da vazão com relação a SST no Estado do Ceará, considerando diferentes defasagens, visando solucionar assuntos relacionados a não linearidade e não normalidade das relações estatísticas, bem como à modificação necessária nos métodos clássicos de análise linear e multivariada. Os preditores selecionados foram o NINO3 e o gradiente da SST do Atlântico.

Tippett *et al.* (2004) simularam um conjunto (*ensemble*) de modelos climáticos de circulação geral (GCM) para estimar a entropia relativa em três regiões onde a SST tropical tem grande impacto na precipitação: Sul da Flórida, Nordeste do Brasil e Quênia. A entropia relativa é uma medida teórica utilizada para quantificar o impacto das variações de SST sobre a precipitação sazonal em relação à variabilidade natural. Eles encontraram que a variação anual da entropia relativa é fortemente correlacionada com as mudanças na precipitação média conjunta e fracamente correlacionada com a variância conjunta. A entropia relativa também é utilizada para ser relacionada com as medidas da capacidade do GCM para reproduzir observações.

Araghinejad *et al.* (2006) apresentaram um método baseado em regressão local para produzir previsões probabilísticas de longo prazo em condições em que não há dados historicamente observados (eventos extremos). O método utiliza o conceito de similaridade entre as observações dos indicadores e conjuntos de observações históricas para obter a melhor estimativa para uma variável dependente. A abordagem apresentada lida com a estrutura de correlação entre as variáveis dependentes e os preditores (SOI, NAO e variáveis hidrológicas) para modelar a incerteza das estimativas. O método proposto foi testado no Zayandeh-rud Rio no Irã. Os resultados mostraram que o método proposto pode melhorar significativamente o desempenho de previsões probabilísticas para uma relação não linear entre preditores hidroclimáticos e vazão.

Voisin *et al.* (2006) descreveram uma série de modelos vinculados à escala regional de demanda de energia elétrica na Califórnia e na região oeste do Pacífico para simular o comportamento das operações dos reservatórios do Rio Columbia, Sacramento e San Joaquin, nos Estados Unidos. Eles identificaram que a demanda de energia elétrica no inverno e na primavera e a produção de energia hidroelétrica anual na região oeste do Pacífico estão relacionadas tanto com o ENOS quanto com o PDO através de variações no clima de inverno.

Maity e Kumar (2008) investigaram a influência de informações de circulação atmosférica de grande escala na variação de vazão na escala da bacia e uma possível melhoria da previsão de vazões, incorporando a informação de tais circulações atmosféricas em grande escala. Foi adotada uma abordagem de rede neural artificial (ANN) usando um algoritmo genético (GA) baseado em otimização evolutiva para modelar a complexa relação entre vazão e circulação atmosférica de grande escala, com os preditores ENOS e Oscilação Equatorial do Oceano Índico.

Tootle *et al.* (2008) aplicaram o método estatístico denominado Decomposição do Valor Singular (SVD) para avaliação das SSTs do Pacífico e do Oceano Atlântico e do regime de vazões em 10 estações hidrométricas da Colômbia. Foi adotada uma abordagem em que as sazonalidades nas SSTs entre primavera-verão (abril a setembro) foram avaliadas para as vazões para o ano-calendário seguinte (janeiro a dezembro). Por meio do índice ENOS, identificou-se que SSTs equatoriais mais quentes (frias) resultaram em menores (maiores) vazões. Identificaram-se regiões SST do Oceano Pacífico (e do Atlântico) e variabilidade hidrológica da Colômbia, especificamente para vazões.

A análise SVD também tem sido usada para identificar as relações entre a variabilidade da SST oceânica e a variabilidade hidrológica. (Tootle at al., 2008) citam diversos trabalhos relacionados, tais como Uvo *et al.* (1998) Wang e Ting (2000) Enfield e Alfaro (2000) e Giannini *et al.* (2000) Rajagopalan *et al.* (2000) Rodriguez-Fonseca e de Castro (2002) Shabbar e Skinner (2004).

Lima e Lall (2008a) testaram o desempenho dos modelos de previsão dos reservatórios do sistema hidroelétrico brasileiro em prever afluências de um a seis meses de antecedência. A informação climática foi incluída nos modelos pelos autores por meio do uso de indicadores climáticos obtidos a partir da temperatura da superfície do mar (SST). A fim de preservar a variabilidade espacial dos dados, foi utilizada a técnica de análise das componentes principais (PCA, em inglês) às séries temporais de afluências.

Em artigo de 2010, Lima e Lall utilizaram índices climáticos em um modelo periódico auto-regressivo exógeno (PARX) para estimar previsões mensais de vazão usando dados de vazão de 54 plantas hidroelétricas no Brasil. Os resultados foram comparados com o modelo clássico de regressão linear, onde o Critério de Informação Bayesiana (BIC, ver Hastie *et al*., 2008) foi utilizado para selecionar o melhor conjunto de preditores a ser incluído no modelo em função do horizonte temporal e o mês de previsão. Comparações com as previsões obtidas com modelos PAR também foram realizadas, para verificar se as informações do clima poderiam melhorar as previsões de vazões pelos modelos em vigor no Brasil à época de seus estudos.

3.7. O PROBLEMA DA INCERTEZA DOS MODELOS

A incerteza sobre a definição do melhor modelo pode ser alta no ramo das previsões. As ferramentas no campo da estatística clássica para lidar com essa incerteza,

além de escassas, podem resultar em conclusões incorretas acerca da definição do melhor modelo, contribuindo para resultados frágeis e viesados.

Ao analisar dados, os pesquisadores deparam-se frequentemente com a definição de quais variáveis deverão ser incluídas nos modelos de previsão. O conjunto apropriado de variáveis explicativas pode ser altamente incerto. Assim, estima-se uma variedade de modelos antes de escolher o modelo de melhor desempenho.

De acordo com Montgomery e Nyhan (2010), a prática da definição do melhor modelo pode resultar em uma série de equívocos. Primeiro, subestima-se a incerteza sobre os efeitos das variáveis de interesse. Baseando-se em inferências sobre um único modelo, assume-se implicitamente que a probabilidade do referido modelo em reproduzir os dados é de 100%, um pressuposto equivocado. Em segundo lugar, alguns pesquisadores costumam ajustar modelos até que a variável explicativa seja estatisticamente significativa, o que pode levar a resultados viesados. E por último, a necessidade em se controlar um grande número de potenciais equívocos pode resultar na superstimação de variáveis, que diminuem a eficiência do modelo sem necessariamente diminuir o viés das variáveis omitidas.

Os métodos clássicos oferecem poucas ferramentas para lidar com a incerteza dos modelos. Os pesquisadores que desejam testar a robustez de suas descobertas muitas vezes estimam uma variedade de modelos para avaliar se a qualidade e/ou a significância dos coeficientes se altera. No entanto, estes testes são realizados de forma aleatória. Além disso, testes de hipóteses frequentistas não oferecem um método para a resolução de conflitos encontrados em todos os modelos.

Outros métodos mais formais para tentar substanciar a seleção dos modelos baseiam-se em métodos estatísticos que penalizam a complexidade dos modelos, como o Critério de Informação de Akaike (AIC) e o Critério de Informação Bayesiano (BIC). Estes métodos partilham uma restrição, a definição do tamanho do potencial modelo. Um modelo com k variáveis independentes implica em 2^k especificações possíveis. A incerteza sobre algumas variáveis de controle faz com que seja extremamente difícil garantir a robustez das especificações dentro de um quadro frequentista. Dado o espaço de modelos relativamente grande, associado a um número módico de variáveis, a incerteza dos modelos pode-se tornar um sério problema (Montgomery e Nyhan, 2010).

A estatístisca clássica não dispõe de técnicas para combinar os resultados dos múltiplos testes de hipóteses em mais medidas de incerteza sobre os coeficientes e / ou

modelos. Uma abordagem bayesiana de ponderação de modelos (BMA, do inglês Bayesian Model Averaging) pode auxiliar os pesquisadores a assegurar que as estimativas dos efeitos das principais variáveis independentes serão robustas para uma ampla gama de possíveis especificações do modelo (Raftery *et al*., 2005).

3.8. A PONDERAÇÃO BAYESIANA DE MODELOS

Entre as diversas técnicas de aprendizado conjunto (*ensemble learning*), destaca-se a ponderação *bayesiana* de modelos (BMA), um método que se utiliza da combinação de distribuições preditivas provenientes de diferentes fontes.

A função densidade de probabilidade (FDP) da previsão BMA de qualquer variável de interesse é uma média ponderada das FDPs, centradas em torno das previsões individuais, onde os pesos para as ponderações são iguais às probabilidades a posteriori dos modelos de geração de previsões e refletem a qualidade dos modelos durante o período de treinamento (período de calibração). Os pesos podem ser utilizados para avaliar a utilidade dos membros do conjunto, que por sua vez podem ser usados como uma base para selecionar esses membros, o que pode ser útil, tendo em vista o tempo de processamento de grandes blocos de dados.

A técnica BMA foi desenvolvida por Leamer (1978), mas ganhou destaque na Estatística em meados da década de 1990, com a divlgação dos estudos de Madigan e Raftery (1994) e Raftery *et al.* (1995) e tem se expandido em áreas como Economia (Fernandez *et al*., 2001), Biologia (Yeung *et al*., 2005), Ecologia (Wintle *et al*., 2003), Saúde Pública (Morales *et al*., 2006), dentre outros. Maiores detalhes podem ser obtidos em Hoeting *et al.* (1999), Hastie *et al.* (2008), Gelman e Hill (2007) e outros.

Na previsão do tempo, a técnica BMA vem sendo amplamente utilizada, sobretudo nos últimos cinco anos, nas previsões por conjuntos (*ensembles* de previsão). Destacam-se os trabalos de Raftery *et al.* (2005), Vrugt *et al.* (2008), Wang *et al.* (2012), Dong *et al*., (2013), Liu e Xie (2014), entre outros.

Na previsão de vazões, citam-se alguns trabalhos, como o de Adamowski *et al.* (2013), Duan *et al.* (2007), Liang e al. (2013), Hemri (2013) e a recente publicação de Lima (2014).

A técnica BMA é um procedimento estatístico que faz inferência nas previsões por meio da ponderação de previsões individuais com base em suas medidas

probabilísticas de verossimilhança, em que as previsões com melhor desempenho recebem pesos maiores do que as previsões de pior desempenho. Além disso, a técnica BMA fornece uma descrição mais confiável da incerteza preditiva total do que o conjunto original, levando a uma FDP mais precisa e melhor calibrada para as previsões probabilísticas (Duan *et al*., 2007).

A mistura de previsões de vários modelos tem o potencial de combinar os pontos fortes dos modelos individuais e melhor representar a incerteza de previsão do que o uso de um único modelo (Wang *et al*., 2012).

Raftery *et al.* (2005) utilizaram a técnica BMA para o pós-processamento de conjuntos (*ensembles)* de dados de previsão do tempo. As FDPs das previsões após a aplicação do método BMA foram melhor calibradas quando comparadas ao *ensemble* de previsões. Os intervalos das previsões BMA foram 62% menores, em média, do que aqueles produzidos com as amostras meteorológicas sem utilizar a técnica de ponderação. Além disso, os dados analisados com a técnica BMA obtiveram o indicador raiz do erro médio quadrático (RMSE) 11% inferior às previsões individuais dos *ensembles*, e 6% menor do que a média dos *ensembles*.

Duan *et al.* (2007) utilizaram nove *ensembles* de previsões hidrológicas para testar e avaliar a aplicação da técnica BMA. Estes conjuntos foram gerados por meio da calibração de três modelos hidrológicos diferentes, utilizando três funções objetivo distintas. Estas funções objetivo foram escolhidas com o intuito de forçar os modelos para que captassem todos os aspectos do hidrograma (altas, médias e baixas vazões). Dois conjuntos de experimentos numéricos foram realizados em três bacias dos EUA para explorar a melhor maneira de usar o método BMA. No primeiro, um único conjunto de pesos BMA foi calculado para se obter previsões BMA, enquanto no segundo, empregaram-se vários conjuntos de pesos, com *ensembles* distintos correspondentes a intervalos de vazão diferentes. Os resultados dos testes mostraram que a técnica BMA tem a vantagem de gerar previsões probabilísticas mais hábeis que as do conjunto original, com a mesma confiabilidade. Os autores concluíram que o desempenho das previsões BMA é geralmente superior ao das melhores previsões individuais. Além disso, concluíram que as previsões BMA, empregando vários conjuntos de pesos, são geralmente melhores do que aquelas que utilizam um único conjunto de pesos.

Wang *et al.* (2012) desenvolveram um modelo BMA de ponderação das previsões de vários modelos estatísticos de previsão de chuvas sazonais ao longo da Austrália usando

índices climáticos como preditores, dando maior peso aos modelos que obtiveram os melhores desempenhos. O estudo apontou para um método BMA capaz de produzir pesos relativamente estáveis na presença de significativa variabilidade na amostragem, resultando em previsões robustas para eventos futuros. As previsões combinaram eficazmente o melhor desempenho dos modelos, representando bem a propagação da incerteza da previsão. O método BMA superou tanto o modelo com dois preditores fixos escolhidos a *priori* como o modelo com base no desempenho preditivo.

Hemri *et al.* (2013) utilizaram a abordagem BMA para quantificar a incerteza na modelagem hidrológica chuva-vazão, utilizando *ensembles* de previsão meteorológica como entrada para seu modelo hidrológico para o rio em Alp, na Suíça, entre 2007 e 2009, abrangendo diferentes horizontes de previsão (*lead times),* desde 1h até 240h. Primeiro, utilizaram a técnica BMA baseada em distribuições gama normais e inversas univariadas, supondo-se a independência entre os diferentes horizontes. Em seguida, os pesquisadores desconsideraram a suposição de independência, a fim de estimar simultaneamente as previsões multivariadas de vazão em toda a faixa de horizontes. A abordagem baseiou-se na ponderação de distribuições normais multivariadas. Tanto as abordagens BMA univariadas como as multivariadas foram capazes de gerar previsões bem calibradas, consideravelmente mais refinadas que a climatologia. Além disso, a configuração BMA multivariada apresentou-se mais vantajosa em relação à abordagem univariada na incorporação de dependências temporais, aumentando a confiabilidade das previsões.

A utilização da técnica BMA para prever combinações entre vários modelos hidrológicos foi o objeto do estudo de Liang *et al.* (2013). Dois modelos hidrológicos foram aplicados ao Rio Amarelo, na bacia do Dongwan, China, para a previsão de cheias. Séries temporais de vazão observadas e simuladas foram transformadas em variáveis normalmente distribuídas e um modelo de mistura gaussiana foi construído por meio da ponderação da distribuição a *posteriori* dos modelos hidrológicos individuais. As probabilidades a *posteriori* das amostras pertencentes a cada modelo hidrológico específico foram adequadas aos respectivos pesos. Os parâmetros do modelo de mistura gaussiana e o peso de cada modelo hidrológico foram estimados pelo algoritmo de "maximização da expectativa" (EM – expectation maximization) e a combinação das previsões na bacia foi obtida a partir dos dois modelos hidrológicos. Em comparação com os modelos unitários, a ponderação BMA ofereceu resultados de previsão de vazão mais precisos. Os autores

concluíram que a técnica BMA é uma abordagem adequada e robusta para a integração entre modelos individuais.

Adamowski *et al.* (2013) estudaram um método de previsão de vazões com base em análises de *wavelets* (ondeletas). Os *ensembles* de previsão foram combinados usando BMA. A abordagem com multi-modelos do tipo *wavelet* ponderados pela técnica BMA foi aplicada em previsões de vazão em diferentes escalas (diária, semanal e mensal) para duas estações nos EUA cujos regimes são muito variados. Verificou-se que simplesmente ponderar os resultados do modelo, atribuindo pesos iguais a cada um, não melhora os resultados em comparação com o melhor modelo unitário. No entanto, a utilização da técnica BMA para a ponderação dos pesos melhorou significativamente o desempenho dos modelos de previsões: entre 9% e 20%, nos modelos mensais; entre 3,85% e 14%, nos modelos semanais; e pequenas melhoras, nas previsões diárias.

Dong *et al.* (2013) efetuaram comparações dos intervalos de confiança obtidos nas previsões geradas pelo método BMA com os de cada modelo individual sob dois esquemas diferentes de combinação BMA. No primeiro esquema BMA, três modelos com a mesma função objetivo (Nash-Sutcliffe) foram calibrados, proporcionando assim três partes de previsões *ensemble* para a combinação BMA. No segundo esquema BMA, todos os três modelos foram calibrados com diferentes funções objetivo (distintas da calibração com base em Nash-Sutcliffe) para obter previsões de nove *ensembles*. Em termos de previsões médias, os *ensembles* BMA do primeiro e do segundo esquema apresentaram desempenho melhor do que qualquer um de seus modelos individuais. O mesmo foi verificado em termos de intervalo de confiança. Também em termos de previsões médias, o *ensemble* BMA do segundo esquema não apresentou vantagem em relação ao BMA do primeiro esquema. O desempenho do BMA do segundo esquema foi melhor do que o do primeiro esquema para as médias e baixas vazões, e pior nas altas. Em termos dos intervalos de confiança, o BMA do segundo esquema apresentou melhores resultados do que o do primeiro esquema, exceção nas maiores vazões.

Liu e Xie (2014) utilizaram a técnica BMA para calibrar os parâmetros de um modelo de previsão quantitativa de probabilidade de precipitação (PQPP) utilizando observações de 43 estações meteorológicas na Bacia de Huaihe (China) e previsões conjuntas de 1 a 7 dias de precipitação acumulada em 24 h. As previsões foram fornecidas por quatro centros de previsão indepententes, cada um com seu sistema de previsão de *ensembles* (EPSs – *Ensemble* Prevision Systems) e por sistemas de previsão de *ensembles*

múltiplos intercambiáveis, fornecidas pelos centros de previsão climática da China. As comparações efetuadas entre os *ensembles* e as técnicas de regressão logística e BMA para os modelos PQPPs indicaram um melhor desempenho dos modelos preditivos BMA em relação às outras previsões.

Lima (2014) utilizou a abordagem BMA para acoplar as previsões de vazões mensais e diárias de 79 anos de dados de vazão afluentes ao reservatório da UHE Itaipu, utilizando um modelo periódico auto-regressivo exógeno (PARX) para as previsões mensais, tendo como preditor exógeno as informações climáticas baseadas em um índice El Niño. As previsões diárias de vazão com horizontes até quatro semanas foram obtidas a partir de um modelo auto-regressivo com coeficientes sazonalmente diferentes. O acoplamento multiescala foi feito por meio do modelo BMA, onde as probabilidades a *posteriori* de cada modelo foram estimadas utilizando amostrador Markov Chain Monte Carlo (MCMC), que responde por incertezas e permite o uso de uma grande variedade de distribuições. Os resultados encontrados indicaram que em horizontes de previsão superiores a 17 dias, o modelo de previsão mensal apresenta maior desempenho para prever as vazões diárias. A técnica BMA possibilitou o acréscimo no indicador de qualidade r^2 em horizontes superiores a 7 dias.

4. REFERENCIAL TEÓRICO

4.1. O PARADIGMA BAYESIANO

O texto a seguir é um resumo elaborado a partir de Ehlers (2014). Referências indicadas para o aprofundamento no tema são Gelman *et al.* (2003), Hastie *et al.* (2008), Robert e Casella (2004).

4.1.1. Teorema de Bayes

Considere uma quantidade de interesse desconhecida θ (tipicamente não observável). A informação de que dispomos sobre θ, resumida através de p(θ), pode ser aumentada observando-se uma quantidade aleatória X relacionada com θ. A distribuição amostral $p(x|\theta)$ define esta relação.

$$p(\theta|x) = \frac{p(x,\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(\theta,x)d\theta} \tag{4.1}$$

Para um valor fixo de x, a função $l(\theta;x) = p(x|\theta)$ fornece a verossimilhança de cada um dos possíveis valores de θ, enquanto p(θ) é chamada distribuição a *priori* de θ. Estas duas fontes de informaçãao, *priori* e verossimilhança, são combinadas levando à distribuição a *posteriori* de θ, $p(\theta|x)$. Assim, a forma usual do teorema de Bayes é

$$p(\theta|x) \propto l(\theta;x)p(\theta) \tag{4.2}$$

(lê-se $p(\theta|x)$ é proporcional a $l(\theta;x)p(\theta)$).

Em outras palavras, temos que a distribuição a *posteriori* é proporcional à verossimilhança x a distribuição a *priori*.

Note que, ao omitir o termo $p(x)$, a igualdade em (4.1) foi substituida por uma proporcionalidade. A partir da forma (4.2) a constante normalizadora da *posteriori* em (4.1) é recuperada como:

$$p(x) = \int p(x, \theta) d\theta = \int p(x|\theta) p(\theta) d\theta = E_\theta[p(X|\theta)] \quad (4.3)$$

que é chamada distribuição preditiva. Esta é a distribuição esperada para a observação x, dado θ.

Suponha então que, após observar $X = x$, estamos interessados na previsão de uma quantidade Y, também relacionada com θ, e descrita probabilisticamente por $p(y|x, \theta)$. A distribuição preditiva de Y dado x é obtida por integração como

$$p(y|x) = \int p(y, \theta|x) d\theta \ = \int p(y|\theta, x) p(\theta|x) d\theta \quad (4.4)$$

Em muitos problemas estatisticos a hipótese de independência condicional entre X e Y dado θ está presente e a distribuição preditiva fica

$$p(y|x) = \int p(y|\theta) p(\theta|x) d\theta \quad (4.5)$$

Note que as previsões são sempre verificáveis uma vez que Y é uma quantidade observável. Finalmente, segue da última equação que

$$p(y|x) = E_{\theta|x}[p(Y|\theta)] \quad (4.6)$$

Os conceitos de *priori* e *posteriori* são relativos àquela observação que está sendo considerada no momento. Assim, $p(\theta|x)$ é a *posteriori* de θ em relação a X (que já foi observado), mas é a *priori* de θ em relação a Y (que não foi observado ainda). Após observar Y= y, uma nova *posteriori* (relativa à $X = x$ e Y = y) é obtida aplicando-se novamente o teorema de Bayes, ou seja, a ordem em que as observações são processadas pelo teorema de Bayes é irrelevante. Na verdade, elas podem até ser processadas em subgrupos.

4.2.PONDERAÇÃO BAYESIANA DE MODELOS

Este item foi baseado em Raftery *et al.* (2005).Além da citada, outras referências no tema são Hastie *et al.* (2008), Vrugt *et al.* (2008), Gelman e Hill (2007), dentre outros.

A Ponderação Bayesiana de Modelos é uma técnica utilizada para a incorporação da incerteza devido à escolha do modelo na inferência estatística.

Considere os K modelos M_1, ... , M_k, tendo como objetivo estimar uma quantidade de interesse Δ e seja D o conjunto de dados para análise. A distribuição a *posteriori* para essa quantidade de interesse é expressa por

$$P(\Delta|D) = \sum_{k=1}^{K} P\,(\Delta|M_k, D)\; P(M_k|D), \tag{4.7}$$

que é uma média da distribuição a *posteriori* sob cada um dos K modelos em consideração, ponderada por suas probabilidades a *posteriori* do modelo respectivo.

A probabilidade a *posteriori* para o modelo M_k, que será o peso w_k do k-ésimo modelo M_k, é dada por

$$P(M_k|D) = \; w_k = \frac{P(D|M_k)P(M_k)}{\sum_{l=1}^{K}\; P(D|M_l)P(M_l)} \tag{4.8}$$

Onde:

$$P(D|M_k) = \int P(D|\theta_k, M_k)P(\theta_k|M_k)d\theta_k \tag{4.9}$$

é a integral da verossimilhança do modelo M_k, θ_k é o vetor de parâmetros do modelo M_k, $P(\theta_k|M_k)$ é a densidade a *priori* de θ_k sob o modelo M_k, $P(D|\theta_k, M_k)$ é a verossimilhança e $P(M_k)$ é a probabilidade a *priori* do modelo M_k ser o verdadeiro modelo. Todas as probabilidades são implicitamente condicionais a M = $\{M_1, \ldots, M_k\}$, conjunto de todos os modelos que estão sendo considerados.

A probabilidade a *posteriori* do modelo M_k irá refletir a capacidade do modelo k em refletir os dados observados. A soma das probabilidades a *posteriori* será igual a um, ou seja, $\sum_{k=1}^{K} w_k = 1$.

Fazendo-se $\hat{\Delta}_k = E[\Delta \mid D, M_k]$, a média e a variância a *posteriori* de Δ podem ser escritas respectivamente por:

$$E[\Delta|D] = \int \Delta \left(\sum_{k=1}^{K} P(\Delta|M_k, D) P(M_k|D) \right) d\Delta$$

$$E[\Delta|D] = \sum_{k=1}^{K} \left(\int \Delta \, P(\Delta|M_k, D) d\Delta \right) P(M_k|D)$$

$$E[\Delta|D] = \sum_{k=1}^{K} \hat{\Delta}_k P(M_k|D) \tag{4.10}$$

$$Var[\Delta| \, D] = \sum_{k=1}^{K} \left(Var[\Delta|D, M_k] + \hat{\Delta}_k{}^2 \right) P(M_k|D) - E[\Delta \mid D]^2 \tag{4.11}$$

Um exemplo simplificado de aplicação da técnica BMA, extraído de Gibbons et al. (2008), encontra-se disponível no Apêndice A. No referido exemplo, considerou-se um par de modelos com o mesmo conjunto de valores de dados, que seguem distribuições normais com médias desconhecidas e possuem apenas um parâmetro cada, gerados por um modelo de referência M_0 que segue distribuição normal padrão (com parâmetros $M_0 \sim N$ [0.0, 1.0]).

4.3. MODELOS AUTO REGRESSIVOS

Séries hidrológicas de intervalo de tempo menor que o intervalo de um ano, tais como séries mensais, têm como característica o comportamento periódico das suas propriedades probabilísticas, como por exemplo, a média, a variância, a assimetria e a estrutura de autocorrelação. A análise deste tipo de séries pode ser feita pelo uso de modelos auto-regressivos periódicos. Estes modelos são referenciados por modelos PAR(p), onde p é a ordem do modelo, ou seja, o número de termos auto-regressivos do modelo (Maceira *et al*., 2006).

Um modelo PAR com média e variâncias periódicas, mas com coeficientes auto-regressivos constantes pode ser representado de acordo com Garcia (2005):

$$z_{\nu,\tau} = \mu_\tau + \sigma_\tau . y_{\nu,\tau} \tag{4.17}$$

e

$$y_{\nu,\tau} = \sum_{j=1}^{p} \phi_j . y_{\nu,\tau-j} + \varepsilon_{\nu,\tau} \tag{4.18}$$

onde $z_{\nu,\tau}$ é uma variável com dependência temporal no ano ν e mês τ, com τ = 1, 2, ..., 12. Assume-se que a variável $z_{\nu,\tau}$ é normalmente distribuída com média μ_τ e variância σ_τ^2 e estrutura de correlação temporal de ordem p, dada por coeficientes auto-regressivos constantes, ϕ_j, conforme apresentado na Equação 4.18.

A variável dependente $y_{\nu,\tau}$ e a variável independente $\varepsilon_{\nu,\tau}$ possuem distribuições normais com média zero e variâncias iguais a 1 e a σ_ε^2, respectivamente. Assim, o conjunto de parâmetros do modelo representado pelas equações 4.17 e 4.18 é $\{\mu_\tau, \sigma_\tau^2, \phi_1, \dots, \phi_p, \sigma_\varepsilon^2;\ \tau = 1,\ 2, \dots, 12\}$.

Considerando a representação do modelo PAR a partir de coeficientes periódicos auto-regressivos, o modelo passa a ser visto como uma forma sintética de representar um conjunto de 12 regressões, uma para cada mês do ano. Em cada uma dessas regressões, assumem-se como variáveis preditoras as vazões aflentes dos p meses anteriores.

A rigor, não há obrigatoriedade em impor para todas as 12 regressões um mesmo número de variáveis preditoras, assim a estrutura de dependência temporal pode apresentar ordem p_τ individual para cada mês. Desse modo, para um mês τ qualquer, tem-se:

$$y_{\nu,\tau} = \sum_{j=1}^{p_\tau} \phi_{j,\tau} . y_{\nu,\tau-j} + \varepsilon_{\nu,\tau} \tag{4.19}$$

Onde $\phi_{j,\tau}$ é o j-ésimo coeficiente auto-regressivo periódico correspondente ao mês τ. A variável independente $\varepsilon_{\nu,\tau}$ passa a apresentar desvios-padrão dos resíduos $\sigma_{\varepsilon\tau}$ periódicos. Nesse caso, o conjunto de parâmetros do modelo passa a ser $\{\mu_t, \sigma_\tau^2, \phi_{1,\tau}, \dots, \phi_{p_\tau,\tau}, \sigma_{\varepsilon\tau}^2;\ \tau = 1,\ 2, \dots, 12\}$.

Maiores informações podem ser obtidas em Box e Jenkins (1970), Salas et al. (1980), Brockwell e Davis (2002).

5. METODOLOGIA

Conforme apresentado, a maioria dos esforços do setor elétrico tem-se concentrado no desenvolvimento de modelos de previsão que utilizam informações meteorológicas a partir de modelos regionais de circulação. A limitada previsibilidade de chuvas num curto prazo resulta em limitadas previsões de vazão para horizontes de tempo superiores a uma semana. O Relatório Anual de Avaliação das Previsões de Vazões, emitidos pelo ONS apresenta os indicadores das previsões para as diversas bacias hidrográficas brasileiras. Como exemplo, as Figuras 5.1 a 5.4 trazem os resultados para a UHE Itaipu, durante o ano de 2013, para a série estimada (correspondendo às previsões da semana corrente) e para as previsões com uma, duas e três semanas de antecedência. Nota-se a perda gradativa de qualidade nas previsões a partir da primeira semana, principalmente no período chuvoso, devido a pouca previsibilidade da precipitação. No período seco, onde a persistência é grande, os modelos funcionam relativamente bem. Essa perda na qualidade no período úmido apresenta-se como potencial oportunidade para melhoria nas previsões.

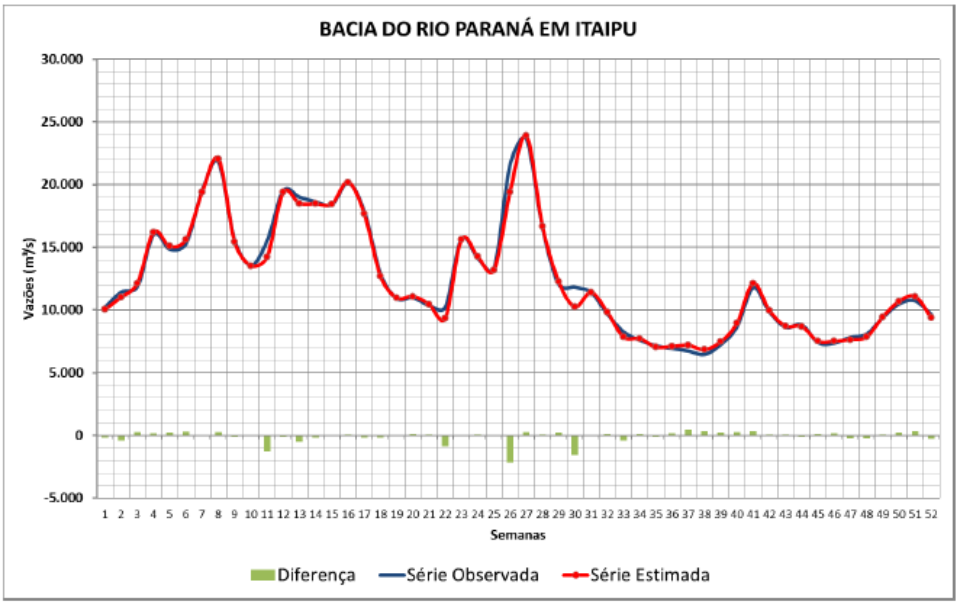

Figura 5.1 - Comparação entre vazão observada e a estimada na semana corrente – bacia do rio Paraná – na UHE Itaipu. Fonte: ONS (2013)

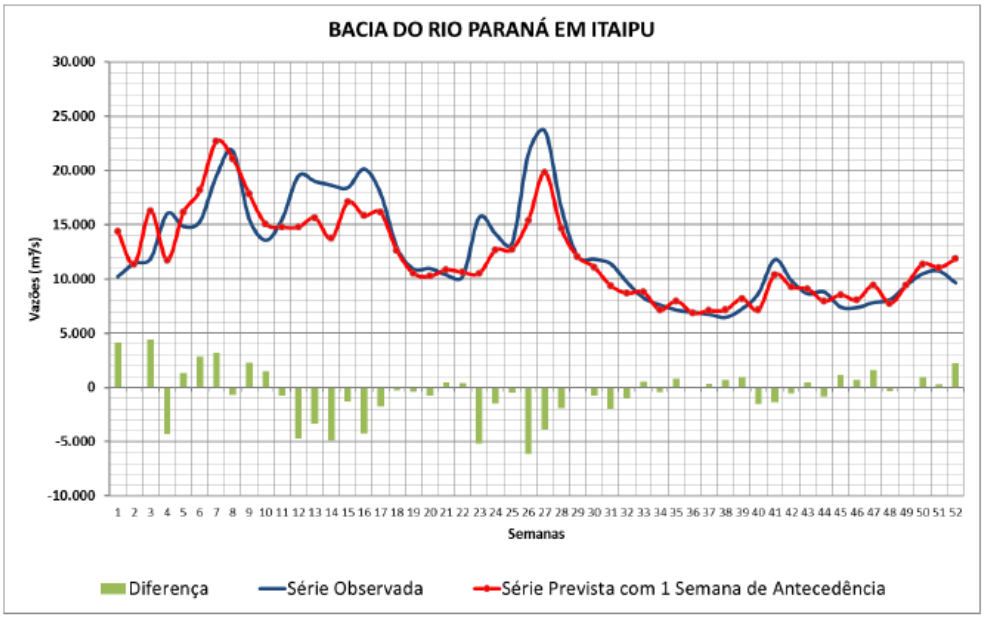

Figura 5.2 - Comparação entre vazão observada e prevista com uma (1) semana de antecedência – bacia do rio Paraná – na UHE Itaipu. Fonte: ONS (2013)

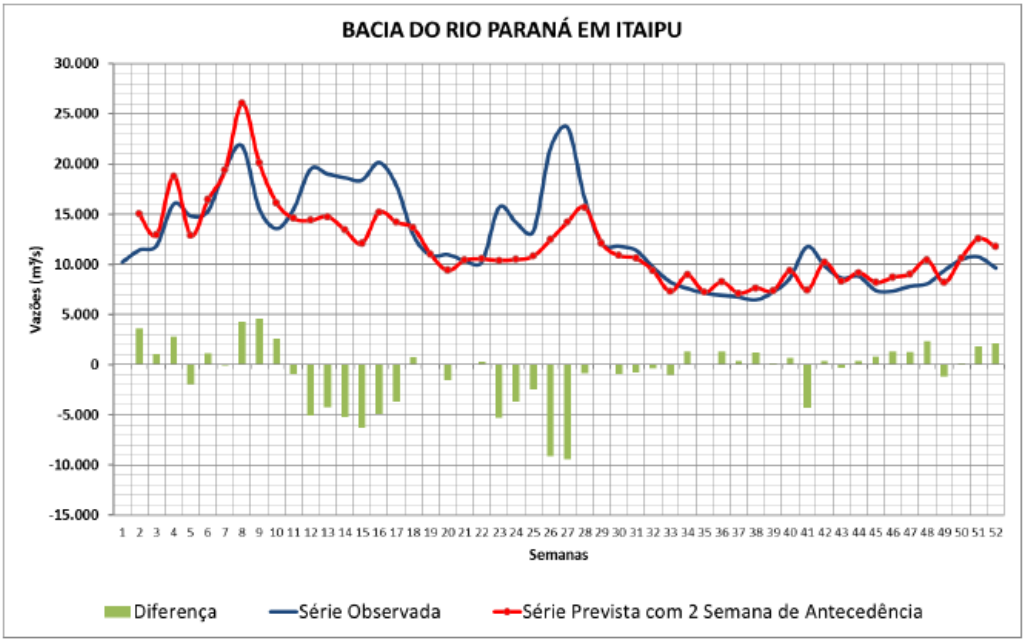

Figura 5.3 - Comparação entre vazão observada e prevista com duas (2) semanas de antecedência – bacia do rio Paraná – na UHE Itaipu. Fonte: ONS (2013)

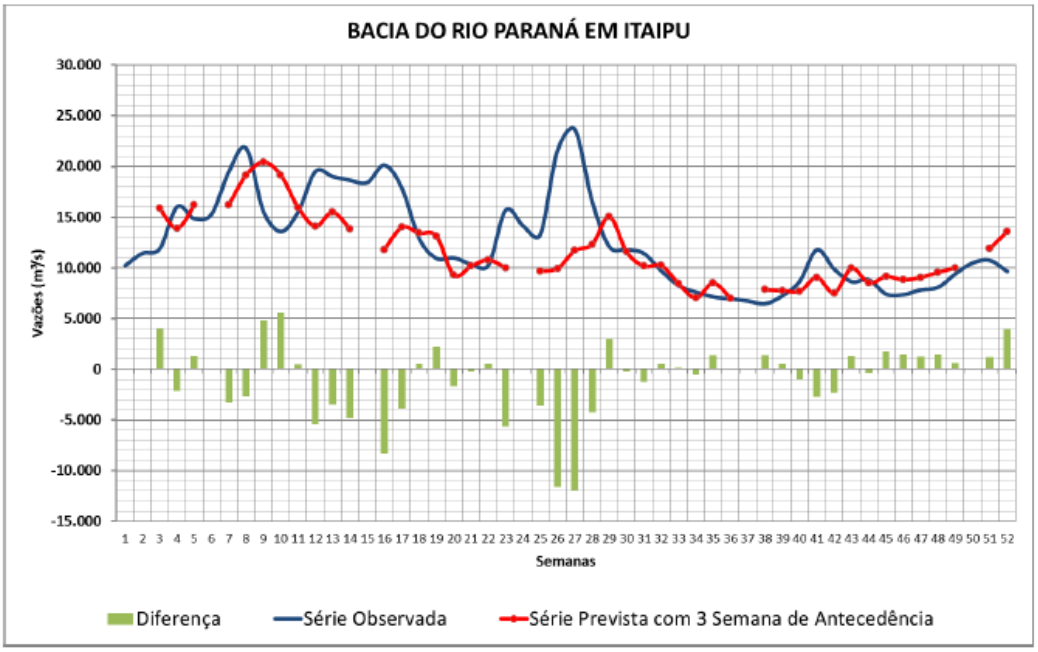

Figura 5.4 - Comparação entre vazão observada e prevista com três (3) semanas de antecedência – bacia do rio Paraná – na UHE Itaipu. Fonte: ONS (2013)

Considerando que as previsões de vazões feitas pelo ONS são feitas principalmente pela classe de modelos auto-regressivos (AR) e não integram informações quantitativas das previsões de vazão de longo prazo, este trabalho busca contribuir com os modelos de previsão de vazões atualmente em uso, utilizando técnicas bayesianas de ponderação de modelos, para que os modelos semanais de previsão utilizados pelo ONS passem a ter uma estrutura hierárquica temporal de previsões, de forma que as vazões geradas para as escalas semanais tenham dependência das vazões obtidas nas escalas mensais, melhorando a qualidade das previsões.

O modelo mensal adotado aqui foi o modelo periódico auto-regressivo exógeno (PARX), estudado por Lima e Lall (2010). O modelo PARX tenta captar a persistência da vazão na parte auto-regressiva e a contribuição do escoamento na parcela exógena através do uso de informações climáticas de larga escala. A utilização de modelos mensais PARX possibilita a associação da variabilidade climática à persistência na variação sazonal da vazão interpretada pelos modelos PAR, diminuindo a variância das previsões e possibilitando o aumento no horizonte de previsões. Com isso, espera-se melhorar as previsões dos modelos semanais utilizados oficialmente no setor elétrico.

5.1. ESTRATÉGIA METODOLÓGICA

A estratégia metodológica para previsão de séries de vazões semanais consistiu de cinco etapas:

1) Adequação do modelo PARX elaborado por Lima e Lall (2010) para a geração de séries mensais para os vinte e oito principais reservatórios de acumulação monitorados pelo ONS, no período compreendido entre janeiro de 2009 e setembro de 2014;

2) Estruturação de um banco de dados contendo os dados de previsão de vazão na escala semanal, provenientes dos relatórios do Programa Mensal de Operação (PMO) emitidos pelo ONS para o mesmo período do modelo mensal (janeiro de 2009 e setembro de 2014);

3) Composição multiescala entre o modelo mensal e os resultados das previsões semanais feitas pelo ONS por meio de um modelo de ponderação bayesiana, com o propósito de melhoria nas previsões semanais.

5.1.1. Modelo mensal

A formulação matemática para o modelo PARX mensal pode ser representada da seguinte forma. Seja q(t+τ) a vazão de um determinado reservatório para o mês t+τ de um certo ano. Os padrões de teleconexões (preditores climáticos) permitem modelar o processo de vazão como uma função de termos persistentes (vazões anteriores) e termos exógenos (preditores climáticos, por exemplo):

$$q(t+\tau) = f(q(t), x_1(t), x_2(t), x_3(t), \dots\dots, x_n(t)) \quad (5.1)$$

onde τ é o horizonte futuro das previsões e $x_1, x_2, x_3, \dots, x_n$ referem-se aos índices climáticos.

Os valores de τ (em meses) podem ser variados de acordo com o horizonte de previsão desejado. Informações climáticas ou hidrológicas anteriores ao tempo t também podem ser incorporadas á Equação 5.1, se necessário.

Uma abordagem comum para modelar a relação expressa na Equação 5.1 é considerar f uma função linear e q(t+τ) uma variável aleatória normalmente distribuída

com preditores e parâmetros que variam sazonalmente, sendo que a variância σ^2 não se mantém constante, pois depende de τ :

$$q(t+\tau) \sim \mathrm{N}\left(\beta_{0t} + \beta_{1t} q(t) + \beta_{2t} x_1(t) + \beta_{3t} x_2(t) + \beta_{4t} x_3(t) + \cdots + \beta_{jt} x_j(t), \sigma_{t,\tau}^2\right) \quad (5.2)$$

Sendo:

q = vazão (afluência);

t= tempo (1 a 12 meses);

j = número de covariáveis utilizadas (persistência e preditores climáticos)

τ = horizonte futuro de previsão, em meses;

β = coeficientes da regressão;

x_1, $x_{2,}$ ...., x_j = covariáveis, que podem ser índices climáticos, informações de precipitação, etc.;

$\sigma_{t,\tau}^2$ = variância, que é função do tempo t e horizonte de previsão τ .

Ao se utilizar índices climáticos como preditores, preserva-se a estrutura espacial de variabilidade dos dados por meio da independência condicional, ou seja, dada a informação climática, as vazões de reservatórios sujeitos as mesmas forçantes climáticas são independentes (Lima e Lall, 2010).

Se os preditores em (5.2) estiverem fortemente correlacionados, então ocorrerá um aumento na variância σ^2 (na incerteza) das previsões, sendo necessária a utilização de alguma técnica de seleção de variáveis para selecionar o subconjunto de indicadores em (5.2) que irá fornecer o menor erro de predição. Adotou-se aqui a técnica de redução de variáveis denominada como *ridge regression* (Hastie *et al*., 2008; Lima e Lall, 2010) .

A técnica de *ridge regression* busca reduzir o número dos coeficientes de regressão, impondo uma penalidade ao seu tamanho. Os coeficientes *ridge* minimizam a soma dos quadrados dos resíduos das penalizações. Os parâmetros de regressão em (5.2) são agora calculados como se segue:

$$\hat{\beta}_t^{\,ridge} = argmin_\beta \left\{ \sum_{i=1}^{N} (q_i(t) - f_i(x,t))^2 + \delta \sum_{j=1}^{4} \beta_{jt}^2 \right\} \quad (5.3)$$

onde δ é um parâmetro adicional (o "parâmetro *rigde*") que determina a redução dos parâmetros β_{jt} (o segundo termo do somatório representa a penalidade imposta aos modelos que utilizam um maior número de parâmetros).

A estimativa ideal para os parâmetros da regressão *ridge* pode ser escrita na forma matricial como:

$$\hat{\beta}_t^{\,ridge} = (X^T X + \delta I)^{-1} X^T q \tag{5.4}$$

onde X é a matriz centrada de entradas e I, a matriz identidade. Note que o problema se reduz a mínimos quadrados ordinários quando δ = 0.

No contexto da análise bayesiana, os parâmetros de *ridge regression* podem ser vistos como uma média a *posteriori*, quando se assume que as distribuições a *priori* para os parâmetros são independentes, distribuições normais com média zero e variância τ^2. Neste caso, $\delta = \frac{\sigma^2}{\tau^2}$, onde σ^2 é dado em (5.2) e assumido conhecido (Hastie *et al*., 2008).

O valor ideal de δ para cada modelo de previsão é obtido depois de avaliar o erro de validação cruzada generalizada (Hastie *et al*., 2008) ao longo de um intervalo de valores para δ e selecionando o valor que produz o menor erro.

5.1.2. Modelo semanal

Foram disponibilizadas pelo ONS as informações de seus relatórios, denominados Programas Mensais de Operação (PMO), que contém os valores verificados e as previsões dos modelos utilizados para cada bacia até seis semanas à frente. Com isso, a etapa de elaboração de um modelo semanal foi suprimida, sendo substituída pelos trabalhos de adequação dos dados, de acordo com a descrição a seguir.

Os dados disponibilizados pelo ONS abrangem as previsões semanais de janeiro de 2009 a setembro de 2014. A Tabela 5.1 traz um exemplo dos dados retirados dos relatórios do PMO, nos meses de agosto e setembro de 2012. A coluna R0 refere-se aos dados do PMO para a UHE Itaipu e contem as previsões de uma a seis semanas operativas à frente. Por convenção, as semanas operativas definidas pelo ONS iniciam-se às 00:00h

de um sábado e terminam às 24:00 da sexta-feira posterior. A primeira semana operativa do mês é aquela que contém o 1º dia do referido mês. Como regra geral, a previsão do PMO é elaborada nas quartas-feiras da última semana operativa do mês e as revisões semanais, nas quintas-feiras da semana operativa anterior.

A cada semana operativa posterior à emissão do relatório R0, o ONS emite um relatório de revisão da semana seguinte, ou seja, a coluna "R1" refere-se à revisão da segunda semana, a coluna "R2" refere-se à revisão da terceira semana e assim sucessivamente. A última linha, denominada "verificado", refere-se às vazões médias semanais verificadas para a UHE Itaipu. Quando o mês tem 5 semanas, como foi o caso de agosto de 2012, são emitidas quatro revisões além da previsão inicial R0; caso contrário, emitem-se três revisões semanais.

Tabela 5.1 - Exemplo das informações oriundas dos relatórios do PMO. Fonte: elaboração própria, adaptado de www.ons.org.br.

				R0	R1	R2	R3	R4	verificado
AGOSTO DE 2012	Semana-1	28/jul	03/ago	9872,0	9192,7	9060,0	8839,0	9065,0	9061,0
	Semana-2	04/ago	10/ago	9608,0	8118,0	7794,4	7666,0	7747,0	7721,0
	Semana-3	11/ago	17/ago	9765,0	7144,0	7084,0	6868,2	7000,0	6971,0
	Semana-4	18/ago	24/ago	8465,0	7130,0	6987,0	6696,0	6439,9	6278,0
	Semana-5	25/ago	31/ago	7824,0	7329,0	7048,0	6856,0	6336,0	5627,2
	Semana-6	01/set	07/set	7456,0	6749,0	6447,0	6415,0	6268,0	5214,0
SETEMBRO DE 2012	Semana-1	01/set	07/set	5280,0	5635,0	5496,0	5610,0		5579,0
	Semana-2	08/set	14/set	5214,0	5361,0	5168,7	5304,0		5289,0
	Semana-3	15/set	21/set	5382,0	5511,0	5143,0	5043,8		5125,0
	Semana-4	22/set	28/set	5564,0	5631,0	5288,0	5906,0		6270,8
	Semana-5	29/set	05/out	6115,0	6281,0	5947,0	6215,0		8973,0
	Semana-6	06/out	12/out	6613,0	6699,0	6275,0	7026,0		9379,0

Assim, como o objetido do trabalho é avaliar os possíveis ganhos que a aplicação da técnica BMA poderá oferecer às previsões para cada semana de antecedência, verifica-se a necessidade de um reposicionamento inicial nos dados disponíveis, de forma a atender o esquema apresentado na tabela a seguir, que ilustra os dados da UHE Itaipu no mês de agosto de 2012.

A cada antecedência semanal, o ONS divulga uma previsão a menos nas revisões do relatório do PMO. Assim, a previsão com seis semanas de antecedência possuirá apenas um valor a cada PMO. Esse é um detalhe inerente à apresentação das informações públicas

nos relatórios do PMO, apesar de o modelo PREVIVAZ estar apto a fornecer previsões até seis semanas a frente. As informações apenas não são divulgadas nos PMOs.

Assim, foi necessário um trabalho de rearranjo dos dados disponíveis, de forma a atender os propósitos deste trabahlo. As Tabelas 5.2 e 5.3 a seguir apresentam a redistribuição dos dados, ordenados pela antecedência das previsões, de uma a seis semanas de antecedência. Assim, as mais de 750 tabelas disponibilizadas pelo ONS foram condensadas em uma única matriz.

Tabela 5.2 – Esquema de reorganização dos dados de acordo com a antecedência semanal. Fonte: elaboração própria, adaptado de www.ons.org.br.

	UHE Itaipu				R0	R1	R2	R3	R4
AGOSTO DE 2012	Semana-1	28/jul	a	03/ago	1sem_anteced				
	Semana-2	04/ago	a	10/ago	2sem_anteced	1sem_anteced			
	Semana-3	11/ago	a	17/ago	3sem_anteced	2sem_anteced	1sem_anteced		
	Semana-4	18/ago	a	24/ago	4sem_anteced	3sem_anteced	2sem_anteced	1sem_anteced	
	Semana-5	25/ago	a	31/ago	5sem_anteced	4sem_anteced	3sem_anteced	2sem_anteced	1sem_anteced
	Semana-6	01/set	a	07/set	6sem_anteced	5sem_anteced	4sem_anteced	3sem_anteced	2sem_anteced

Tabela 5.3 – Dados da Tabela 5.1, ordenados de acordo com a antecedência das previsões semanais. Fonte: elaboração própria, adaptado de www.ons.org.br.

				1s_antec	2s_antec	3s_antec	4s_antec	5s_antec	6s_antec		verificado
AGOSTO DE 2012	Semana-1	28/jul	03/ago	9872,0							9061,0
	Semana-2	04/ago	10/ago	8118,0	9608,0						7721,0
	Semana-3	11/ago	17/ago	7084,0	7144,0	9765,0					6971,0
	Semana-4	18/ago	24/ago	6696,0	6987,0	7130,0	8465,0				6278,0
	Semana-5	25/ago	31/ago	6336,0	6856,0	7048,0	7329,0	7824,0			5627,2
	Semana-6	01/set	07/set		6268,0	6415,0	6447,0	6749,0	7456,0		5214,0
SETEMBRO DE 2012	Semana-1	01/set	07/set	5280,0							5579,0
	Semana-2	08/set	14/set	5361,0	5214,0						5289,0
	Semana-3	15/set	21/set	5143,0	5511,0	5382,0					5125,0
	Semana-4	22/set	28/set	5906,0	5288,0	5631,0	5564,0				6270,8
	Semana-5	29/set	05/out		6215,0	5947,0	6281,0	6115,0			8973,0
	Semana-6	06/out	12/out			7026,0	6275,0	6699,0	6613,0		9379,0

Após a redistribuição dos dados, de acordo com as Tabelas 5.2 e 5.3, foram acrescentadas duas colunas às matrizes de dados, contendo os resultados das previsões resultantes do modelo mensal PARX, com um e dois meses de antecedência, devidamente convertidos para vazões médias semanais. Genericamente, a matriz resultante tem a estrutura ilustrada pela Tabela 5.4.

Tabela 5.4 – Visão esquemática da matriz de dados primários para a composição BMA.

Semana operativa do ONS	Vazão Verificada (Y)	Previsões semanais do ONS				Previsões do Modelo Mensal	
		1 semana anteced. (X_1)	2 sem. anteced. (X_2)	.....	6 sem. anteced. (X_6)	1 mês anteced. (X_7)	2 meses anteced. (X_8)
Sem1.Jan.2009	y_1	x_{11}	x_{21}	...	x_{61}	x_{71}	x_{81}
Sem2.Jan.2009	y_2	x_{12}	x_{22}	...	x_{62}	x_{72}	x_{82}
Sem3.Jan.2009	y_3	x_{13}	x_{23}	...	x_{63}	x_{73}	x_{83}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Sem6.Set.2014	y_n	x_{1n}	x_{2n}	x_{7n}	x_{6n}	x_{7n}	x_{8n}

A matriz genérica foi então dividida em seis sub-matrizes, uma para cada previsão semanal, cada sub-matriz contendo 3 colunas. A primeira sub-matriz, referente às previsões semanais realizadas pelo ONS com uma semana de antecedência, foi composta do vetor Y (valores observados), do vetor X_1, com as previsões semanais conduzidas pelo ONS com uma semana de antecedência, e do vetor X_7, com as previsões do modelo mensal PARX. A segunda matriz, referente às previsões semanais do ONS com duas semanas de antecedência, foi composta do vetor Y, do vetor X_2 e do vetor X_7. Para as demais semanas, o mesmo método foi repetido, sendo que o vetor X_8 compôs a terceira coluna da sub-matriz da sexta semana. Para a sub-matriz da quinta semana, foi elaborado um algoritmo que seleciona, para a terceira coluna, ou valores do vetor X_7, quando o mês operativo possui 5 semanas, ou valores do vetor X_8, caso contrário.

5.1.3. Ponderação dos modelos mensal e semanal

A idéia básica da aplicação da técnica BMA é melhorar as previsões de vazões semanais, sobretudo em horizontes de previsão que vão além da persistência da vazão (que dura em torno de dez dias, ou seja, pouco mais de uma semana), agregando as informações fornecidas pelas previsões mensais, que, na ausência do modelo de previsão semanal,

fornecerá as melhores previsões semanais (cada semana do mês recebe o mesmo valor da previsão feita mensalmente).

Uma forma possível de realizar tal tarefa para cada reservatório é a ponderação das previsões feitas pelos modelos semanal e mensal de acordo com alguns pesos pré-definidos w como um modelo de mistura finita (Lima, 2014):

$$q^*(t) = w_1 q(t) + w_2 Q(t) + \epsilon_\tau^*(t) \tag{5.5}$$

Onde:

$q^*(t)$: previsão de vazão semanal após a ponderação bayesiana;

t: antecedência semanal (lead time), no nosso caso t = 1 a 6;

$q(t)$: vazão prevista mensal (saída do modelo PARX);

$Q(t)$: vazão prevista semanal (saída dos modelos do ONS);

$\epsilon_\tau^*(t)$: desvios das previsões.

Aqui, assume-se que $\epsilon_\tau^*(\mathrm{t})$ tem distribuição normal com média zero e variância $\tau_{\epsilon_\tau}^2(\mathrm{t})$. Os termos w_1 e w_2 representam os pesos bayesianos dos modelos mensal e semanal, respectivamente, cuja soma $w_1 + w_2 = 1$;

Em um modelo de ponderação bayesiana, a Eq. (5.5) pode ser reescrita como:

$$P_r(Predição|Dados) = \sum_{i=1}^{2} P_r(Predição|Dados, M_i).\, P_r\,(M_i|Dados), \tag{5.6}$$

onde o termo Predição refere-se à previsão semanal em alguma semana t , M_1 refere-se às previsões semanais obtidas dos relatórios do PMO, M_2 é a previsão mensal (Equação 5.2) e $P_r(\mathrm{Mi} \mid Dados) = w_i$ é a probabilidade a *posteriori* para modelo M_i (peso de cada modelo), dada por:

$$P_r(\mathrm{M}_i \mid Dados) = \frac{P_r(Dados|M_i)\; P_r(M_i)}{\sum_{k=1}^{2} P_r(Dados|M_k)\; P_r(M_k)}, \tag{5.7}$$

onde $P_r(Dados \mid M_i)$ é a verossimilhança do modelo i e P_r (M_k) corresponde à probabilidade a *priori* de que M_k é o modelo verdadeiro.

Como foram adotadas as saídas dos modelos semanais utilizados pelo ONS como insumo do modelo de ponderações, ou seja, dados do tipo determinísticos, o termo $P_r(\text{Predição}|\text{Dados}, M_1) = 1$. Para o modelo mensal, assume-se que:

$$P_r(Predição|Dados, M_2) =$$
$$= N(q(t)|\beta_0 + \beta_{1t}.Q_{t-1} + \beta_{2t}x_1(t-1) + \beta_{3t}x_2(t-1) + \beta_{4t}x_3(t-1), \sigma_t^2) \quad (5.8)$$

Onde $\sigma_t^2 = 0$, com os demais parâmetros associados sendo estimados de forma independente, utilizando dados históricos para cada modelo.

Os pesos desconhecidos w foram estimados utilizando o algoritmo de otimização *expectation maximization* (EM), cuja fundamentação encontra-se no Apêndice 2 desta dissertação, mas poderiam ser adotadas outras abordagens, por exemplo, técnicas de amostragem do tipo Monte Carlo via cadeias de Markov, adotada por Lima (2014).

5.1.4. Métricas de qualidade adotadas

Nas análises efetuadas nesta dissertação, foram adotados quatro indicadores de qualidade, a saber: a raiz do erro médio quadrático (RMSE), o erro médio percentual absoluto (MAPE), o coeficiente de eficiência de Nash-Sutcliffe (NS) e o indicador distância multicritério (DM). Os três últimos indicadores foram escolhidos por serem os mesmos adotados pelo ONS em suas avaliações de qualidade das previsões. Segundo informações retiradas dos Relatórios Anuais de Previsão de Vazões do ONS, a adoção do indicador MAPE é indicada para avaliar a qualidade de "curto prazo" das previsões. Já a qualidade de "longo prazo", é mais bem traduzida pelo indicador NS.

Os Relatórios Anuais de Previsão de Vazões apresentam as formulações de cada indicador e suas interpretações.

Erro Médio Percentual Absoluto – MAPE:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{(O_t - P_t)}{O_t}\right| \quad (5.9)$$

Coeficiente de Eficiência de Nash-Sutcliffe – NS

$$NS = 1 - \frac{\sum_{t=1}^{n}(P_t - O_t)^2}{\sum_{t=1}^{n}(O_t - \bar{O})^2} \quad (5.10)$$

Onde:

n - número de intervalos de tempo

O_t - vazão observada no intervalo de tempo t

$\bar{O}$- média das vazões observadas nos intervalos de tempo de 1 a n

P_t - vazão prevista no intervalo de tempo t;

O MAPE é a média da diferença absoluta entre os valores previstos e observados, expressa em percentagem dos valores observados. Segundo o ONS, este indicador apresenta a vantagem de não ser influenciado pelo valor absoluto das vazões. Este indicador pode assumir valores no intervalo $(\infty,0]$, sendo MAPE=0 o valor correspondente a previsões perfeitas no período.

O NS é utilizado para avaliar o poder de predição dos modelos, podendo assumir valores no intervalo $(-\infty,1]$, sendo que NS=1 corresponde à previsão perfeita das vazões no período em foco. Quando NS<0, entende-se que utilizar as previsões do modelo é pior do que utilizar o valor médio das observações.

O índice Distância Multicritério – DM, como pode ser visto na Figura 5.5, utiliza os indicadores MAPE e NS como um par ordenado (1-NS, MAPE), e calcula a distância euclidiana ao ponto ótimo, no caso (0,0). A equação para o indicador DM é:

$$DM = \sqrt{(1 - NS)^2 + MAPE^2} \quad (5.11)$$

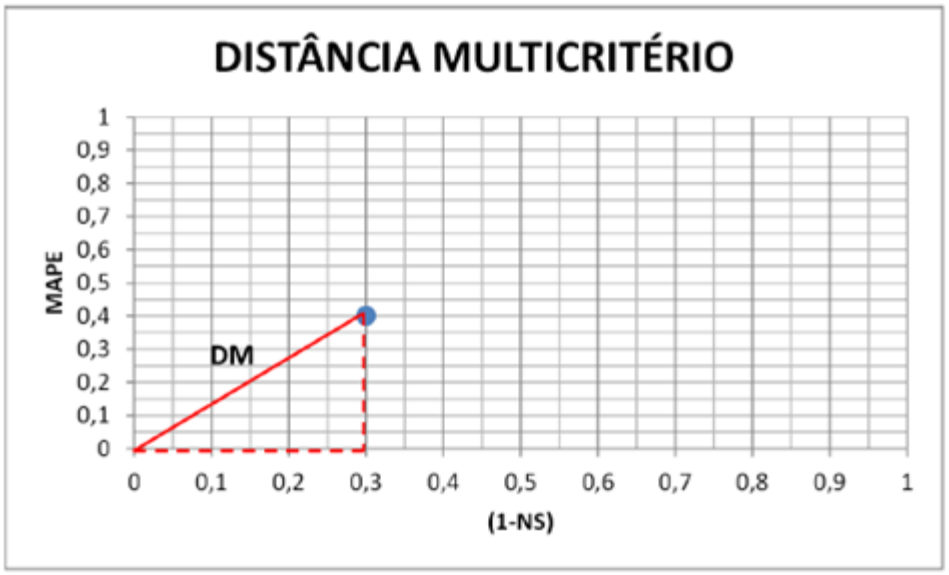

Figura 5.5 - Representação gráfica da distância multicritério. Fonte: www.ons.org.br.

5.1.5. O teste estatístico de Diebold-Mariano

No intuito de verificar se as diferenças entre as competências dos modelos semanais devem-se a erros de amostragem, foi testada a hipótese nula de que as saídas dos modelos semanais antes e após a ponderação BMA possuem iguais acurácias de precisão, utilizando a estatística de teste Diebold-Mariano (Diebold e Mariano, 1995). Para amostras com tamanho moderadamente grandes, o teste de Diebold-Mariano é robusto quando os resíduos dos dois métodos estão correlacionados, aplicando-se a uma grande variedade de distribuições dos resíduos.

Considerando que $\{y_t\}$ denote a série de vazões observada e que $\{\hat{y}_{i,t}^h\}$ denote as saídas dos modelos de previsões com antecedência semanal h = 1 a 6, sendo que i=1 representa as saídas dos modelos do ONS e i=2, as saídas do modelo BMA. Supondo que os erros de previsão dos modelos concorrentes sejam representados por:

$$e_{i,t}^h = y_t^h - \hat{y}_{i,t}^h \tag{5.12}$$

A precisão de cada previsão é medida por funções de perda do tipo:

$$L(y_t^h, \hat{y}_{i,t}^h) = L\left(e_{i,t}^h\right) \tag{5.13}$$

Existem uma série de funções de perda, sendo a função de erros absolutos e a função de erros quadráticos as mais usuais. Ambas as funções são simétricas em torno do ponto de origem. A seguir, apresenta-se a estrutura das duas referidas funções de perda, sendo que nesta dissertação adotou-se a função quadrática (L_2), que penaliza os maiores erros de forma mais acentuada.

Função de perda de erros absolutos:

$$L_1(y_t^h, \hat{y}_{i,t}^h) = L_1(e_{i,t}^h) = \sum_{t=1}^{T} |e_{i,t}^h|, \quad com\ h = 1, \dots, 6\ (no\ presente\ caso) \qquad (5.14)$$

Função de perda de erros quadráticos:

$$L_2(y_t^h, \hat{y}_{i,t}^h) = L_2(e_{i,t}^h) = \sum_{t=1}^{T} (e_{i,t}^h)^2, \quad com\ h = 1, \dots, 6\ (no\ presente\ caso) \qquad (5.15)$$

Para determinar se a saída do modelo BMA (modelo 2) possui maior acurácia nas previsões do que a saída dos modelos do ONS (modelo 1), foi testada a hipótese nula de que os dois modelos possuem a mesma acurácia. Omitindo o índice h referente às antecedências semanais, a hipótese nula pode ser representada como:

$$H_0 = E[L(e_{1,t})] = E[L(e_{2,t})] \qquad (5.16)$$

A hipótese alternativa, considerando que o modelo BMA possui maior acurácia nas previsões é por:

$$H_1 = E[L(e_{1,t})] > E[L(e_{2,t})] \qquad (5.17)$$

O teste de Diebold-Mariano é baseado na diferença d_t das funções de perda:

$$d_t = L(e_{1,t}) - L(e_{2,t}) \qquad (5.18)$$

De forma equivalente, a hipótese nula de igual precisão nas previsões é definida como $H_0 = E[d_t] = 0$. Adotando-se a média amostral da perda diferencial $\bar{d}$ como sendo:

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t = \frac{1}{T}\sum_{t=1}^{T}\left[L(e_{1,t}) - L(e_{2,t})\right] \tag{5.19}$$

a estatística do teste de Diebold-Mariano (DM_{test}) pode ser escrita:

$$DM_{test} = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}} \xrightarrow{d} N(0,1) \tag{5.20}$$

onde $2\pi\hat{f}_d(0)$ é um estimador consistente da variância assintótica de $\sqrt{T}\bar{d}$.

A estatística de teste baseia-se na média e no desvio da diferença entre o erro quadrático da previsão dos dois modelos concorrentes, em que a estimativa da variância representa a correlação do erro de previsão de acordo com a antecedência da previsão (reduzindo em seguida, os graus de liberdade). Diebold e Mariano (1995) mostraram que a distribuição da estatística DM_{test} converge para uma distribuição normal padrão.

5.1.6. O pacote computacional *ensembleBMA*

A linguagem de programação R (R Core Team, 2014) permite a utilização de pacotes de programação disponíveis para o estudo da abordagem de Ponderação Bayesiana de Modelos (BMA). Podemos citar os pacotes BMA – versão 3.17.1 (Raftery *et al*., 2014) e BMS – versão 0.3.0 (Feldkircher e Zeugner, 2009). Nesta dissertação, foi utilizado o pacote *ensembleBMA*, versão 4.5, desenvolvido por Fraley *et al.* (2010). Trata-se de uma série de funções probabilísticas para previsões meteorológicas, que permitem o pós-processamento de *ensembles* (conjuntos) por meio da aplicação da técnica BMA.

Os modelos de previsão *ensembleBMA* usam misturas de distribuições, onde cada componente corresponde a um membro do conjunto de previsões, e a forma da distribuição do componente depende do parâmetro climático em questão (temperatura, precipitação, velocidade do vento, por exemplo). Os parâmetros do modelo são estimados por meio do algoritmo de otimização EM, a partir de um conjunto de dados de treinamento, definido pelo usuário. O pacote conta atualmente com 3 opções de ajustes de modelos: mistura de distribuições normais (gaussianas), de distribuições Gama e de distribuições Bernoulli-

Gama. O pacote *ensembleBMA* inclui funções para a geração de novas previsões e para a avaliação do desempenho preditivo do modelo ajustado.

A abordagem BMA para o pós-processamento de conjuntos de previsões foi introduzida por Raftery *et al.* (2005) e os detalhes podem ser encontrados em Gneiting e Raftery (2007), Gneiting *et al.* (2007).

6. RESULTADOS E DISCUSSÕES

6.1. RESULTADOS PARA AS MODELAGENS MENSAIS

O modelo mensal foi elaborado considerando o conjunto dos vinte e oito principais reservatórios de acumulação monitorados pelo ONS, cujo acompanhamento semanal pode ser efetuado por meio do acesso ao boletim "Situação dos Principais Reservatórios de Acumulação", na página eletrônica do ONS. A Figura 6.1 apresenta a localização desses reservatórios, sendo que a divisão das regiões geográficas no Brasil segue a padronização sugerida por Grimm (2004). As principais bacias hidrográficas que contém as hidroelétricas analisadas estão destacadas em vermelho e numeradas de 1 a 6, sendo: 1 - Bacia do Tocantins, 2 - Bacia do São Francisco; 3 - Bacia do Atlântico Leste; 4 - Bacia do Paraná; 5 - Bacia do Uruguai e 6 - Bacia do Atlântico Sudeste. Maiores detalhes acerca da localização dos vinte e oito reservatórios estudados podem ser consultados na Tabela 6.1.

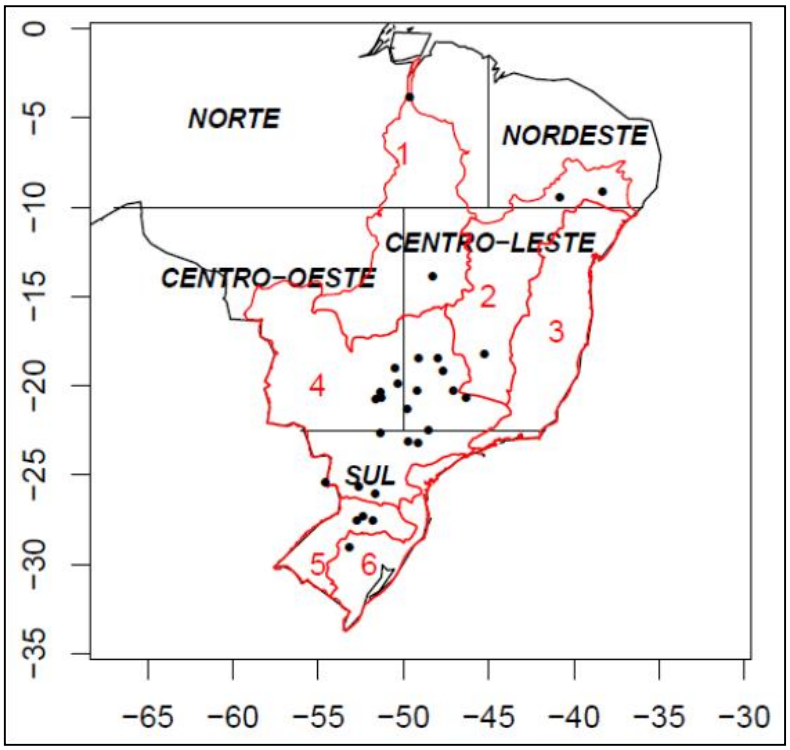

Figura 6.1 - Localização dos sítios hidroelétricos utilizados na modelagem. As regiões geográficas (linhas pretas) são definidas como em Grimm (2004). As linhas vermelhas mostram a delimitação geográfica das bacias hidrográficas que contém as hidroelétricas analisadas aqui. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

Tabela 6.1 - Identificação dos 28 reservatórios que fazem o escopo deste estudo, sua localização (rio), a área de drenagem aproximada e o(s) modelo(s) de previsões semanais utilizados pelo ONS. Fonte: elaboração própria, adaptado de www.ons.org.br.

Reservatório / Área de drenagem	**Rio**	**Modelo Hidrológico (ONS)**
Furnas / 50.464 km^2	Grande	SMAP Previvaz
Mascarenhas de Moraes / 59.730 km^2		
Marimbondo / 118.515 km^2		
Água Vermelha / 139.437 km^2		
Emborcação / 29.050 km^2	Paranaíba	
Nova Ponte / 15.480 km^2	Araguari (afluente Paranaíba)	
Itumbiara / 94.728 km^2	Paranaíba	
São Simão / 171.474 km^2		MGB Previvaz
Jurumirim (A.A. Laydner) / 5.768 km^2	Paranapanema	SMAP Previvaz
Chavantes / 27.769 km^2		
Capivara / 84.715 km^2		
Foz do Areia (G.B. Munhoz) / 30.100 km^2	Iguaçu	Fuzzy Previvaz
Salto Santiago / 43.330 km^2		
Machadinho / 32.050 km^2	Uruguai	MPCV Previvaz
Itá / 44.500 km^2	Uruguai	
Passo Fundo / 2.300 km^2	Passo Fundo (afluente Uruguai)	
Passo Real / 8.140 km^2	Jacuí	
Três Marias / 50.732 km^2	São Francisco	NEURO3M Previvaz
Sobradinho / 499.084 km^2		CPINS NeuroSF Previvaz
Itaparica / 593.384 km^2		
Barra Bonita / 33.156 km^2	Tietê	Previvaz
Promissão / 58.106 km^2		
Três Irmãos / 71.221 km^2		
Ilha Solteira / 377.197 km^2	Tietê/Paraná	
Jupiá / 476.797 km^2	Paraná	MGB SMAP Previvaz
Itaipu / 823.555 km^2		
Serra da Mesa / 51.233 km^2	Tocantins	Previvaz
Tucuruí / 757.577 km^2		

Para fins de compatibilização, tanto as séries médias mensais de vazão dos vinte e oito reservatórios quanto às séries mensais dos preditores climáticos possuem período idêntico, e vão de janeiro de 1949 a setembro de 2014. As séries semanais iniciam em janeiro de 2009, com término também em setembro de 2014.

As séries médias mensais históricas de vazão naturais, disponíveis em http://www.ons.com.br/operacao/vazoes_naturais.aspx, foram adotadas entre janeiro de 1949 até dezembro de 2008. O período entre janeiro de 1931 a dezembro de 1948 não foi utilizado, por não haver dados referentes aos preditores climáticos nesse período. Entre janeiro de 2009 e setembro de 2014, os valores das séries mensais foram compatibilizados com as informações das séries semanais, disponíveis nos relatórios do PMO.

Os dados interpolados das anomalias de temperatura da superfície do mar (SST) provenientes dos oceanos Tropical, Atlântico e Pacífico (Kaplan *et al*., 1998; Reynolds e Smith, 1994) são fornecidos pelo International Research Institute for Climate and Society (IRI) em http:// iridl.ldeo.columbia.edu/SOURCES/.KAPLAN/.EXTENDED/.v2/.ssta/.

Os dados do vento zonal de baixa pressão (700 mb) são disponibilizados em http://www.iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEPNCAR/.CDAS/.MONTHLY/.Intrinsic/.PressureLevel/.u/ pelo NOAA NCEP-NCAR.

Os dados dos preditores climáticos e de vazão foram primeiramente divididos em um conjunto de calibração (abrangendo o período de 1949-2008), onde os parâmetros do modelo são estimados, e um conjunto de validação (2009-2014), que avalia o desempenho do modelo. Um comparativo entre o desempenho do modelo ao se utilizar o período de clibração completo e o de diversos períodos de calibração contendo dez anos cada foi efetuado, confirmando-se pela opção de escolha do período completo, que alcançou melhores resultados, como pode ser observado na Figura 6.2 e na Tabela 6.2.

O longo período de calibração adotado visa garantir que todos os reservatórios tenham passado por uma variedade de condições de vazão (tanto em condições normais, como abaixo e acima das médias). Assim, ao experimentar as diversas transições de períodos secos e úmidos, espera-se que o período de calibração escolhido minimize a dependência das competências do modelo e das estimativas de parâmetros.

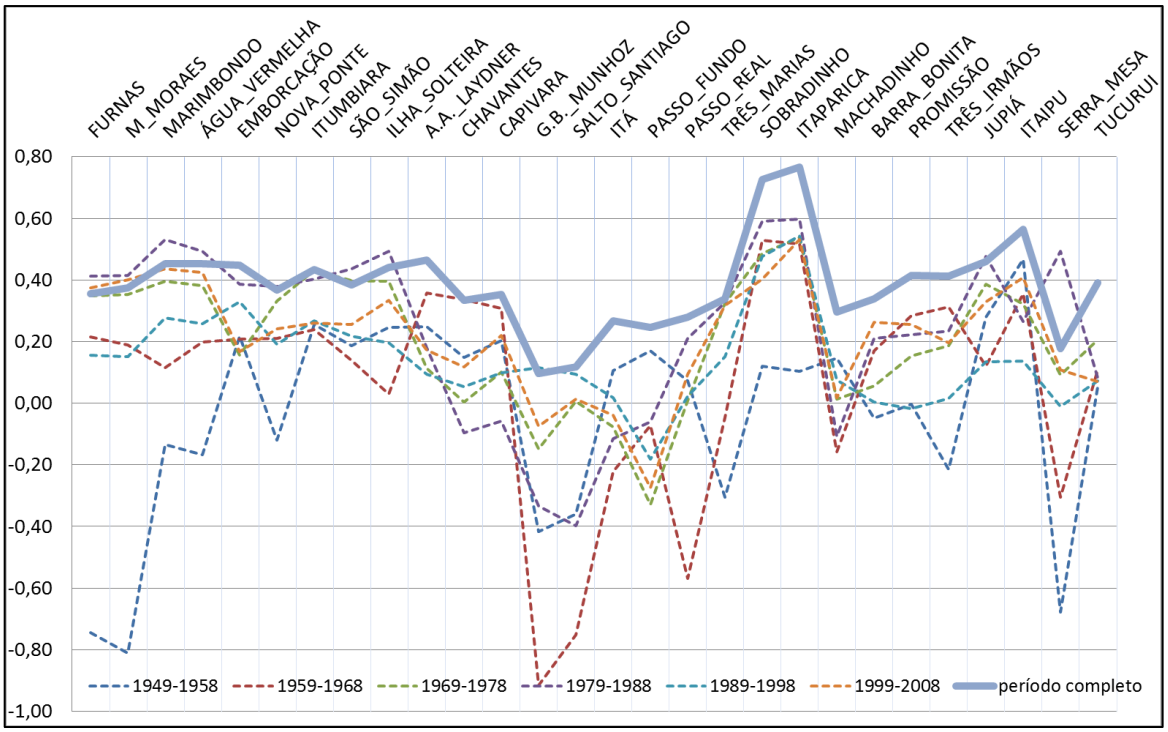

Figura 6.2 – Comparativo do indicador NS entre diversos períodos de calibração para o modelo PARX mensal.

Tabela 6.2 – Valores médios do indicador NS para cada período de calibração ilustrado na Figura 6.2.

1949-1958	-0,04
1959-1968	0,07
1969-1978	0,20
1979-1988	0,24
1989-1998	0,14
1999-2008	0,22
Período completo (1949 a 2008)	0,38

6.1.1. Índices Climáticos

As chuvas e os padrões de vazão em todo o Brasil são as mais afetadas pela SST a partir de três regiões distintas (Lima e Lall, 2010):

(i) Oceano Pacífico tropical, associado aos eventos El Niño e La Niña;

(ii) Oceano Atlântico subtropical, acompanhado por mudanças na Zona de Convergência do Atlântico Sul (ZCAS); e

(iii) Oceano Atlântico tropical, combinado com o deslocamento da Zona de Convergência Intertropical (ZCIT).

Para os propósitos deste trabalho, as duas primeiras regiões são mais importantes, já que a última é mais associada com mudanças nos padrões pluviométricos do Nordeste do Brasil e da região amazônica, onde poucos reservatórios hidroelétricos estão atualmente instalados. Com a crescente instalação de hidroelétricas na Amazônia, trabalhos futuros poderão incorporar esse fator nos estudos de vazão desses futuros reservatórios.

Após a análise da correlação entre as séries de vazão dos reservatórios e alguns índices climáticos relacionados às duas primeiras teleconexões, optou-se por considerar 3 índices climáticos, descritos a seguir. A Figura 6.3 mostra a localização espacial dos índices adotados nesse trabaho.

- Índice NINO3, definido como a anomalia na temperatura média mensal da superfície do mar (SST) calculada sobre a área geográfica com 5°N – 5°S de latitude e 150°W – 90°W de longitude,
- Índice SST, obtido pela média da anomalia de SST sobre a região de 12°S – 30°S de latitude e 40°W – 20°W de longitude.
- Índice U1, resultado da média dos valores do vento zonal de baixa pressão (700 mb) sobre a região de 10°S – 20°S de latitude e 50°W – 35°W de longitude.

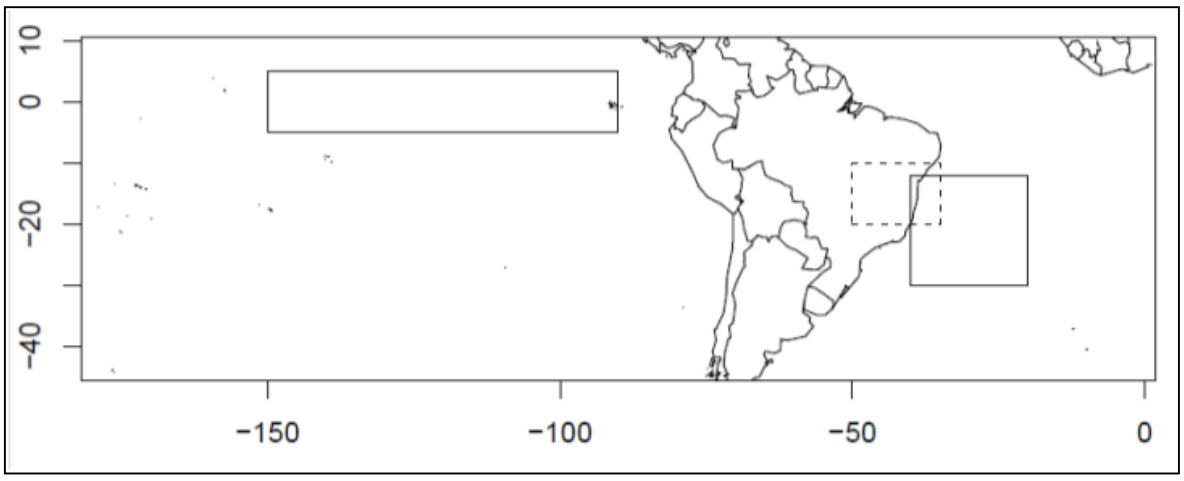

Figura 6.3 - Localização espacial dos índices climáticos baseados na SST (retângulos com linha sólida, sendo que o mais à esquerda se refere ao índice NINO3) e no vento zonal de baixa pressão (retângulo com linha tracejada). Fonte: Lima e Lall, 2010.

A Figura 6.4 mostra as correlações entre o índice NINO3 e os dados de vazão para março com defasagens *(lags)* de um a dois meses. Os círculos em azul (vermelho) apresentam os resultados em que as correlações são negativas (positivas). As mesmas informações são apresentadas para o mês de setembro, na Figura 6.5.

Para os reservatórios localizados nas regiões Sul e Sudeste do Brasil, março e setembro são os meses abaixo e acima da vazão média, respectivamente (Lima e Lall, 2010). As correlações positivas observadas em ambos os meses indicam que, durante os eventos de El Niño, condições mais úmidas prevalecem tanto no período seco como no período chuvoso. Os resultados encontrados acompanham àqueles de Lima *et al.* (2007), que concluíram que os picos de vazão da estação chuvosa e as vazões mínimas da temporada seca da maioria dos reservatórios hidroelétricos no Brasil (exceto aqueles localizados no sul do Brasil) ocorrem em média, durante março e setembro.

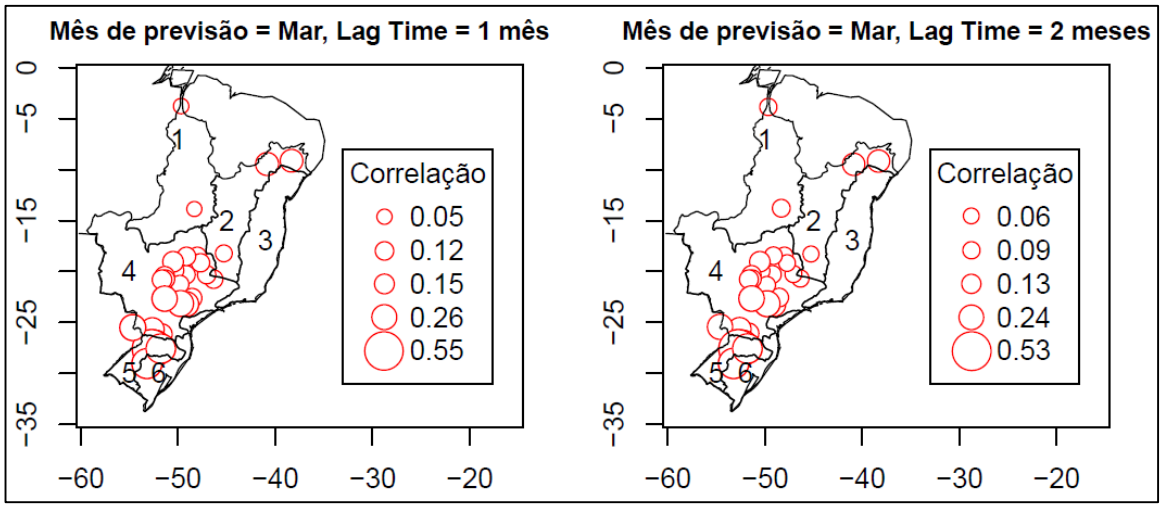

Figura 6.4 - Correlações entre as vazões do mês de março e o índice NINO3, com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

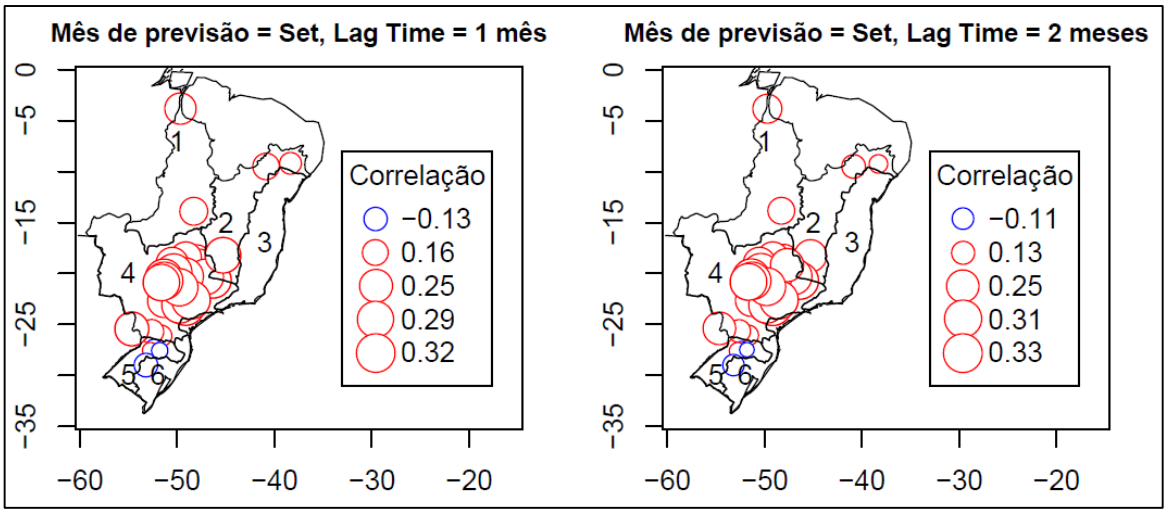

Figura 6.5 - Correlações entre as vazões do mês de setembro e o índice NINO3, com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

A Figura 6.6 apresenta os mesmos resultados das duas figuras anteriores, só que desta vez para os valores das correlações médias anuais, indicando que, na média, prevalece a relação positiva entre as vazões e o preditor NINO3.

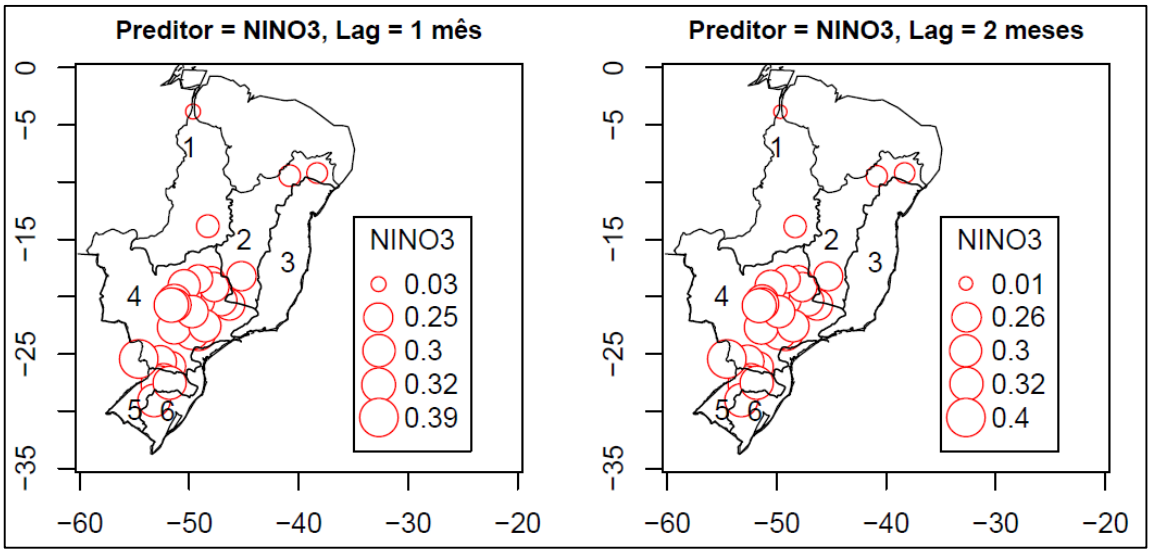

Figura 6.6 - Valores médios anuais das correlações entre as vazões e o índice NINO3, com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

A Figura 6.7 mostra padrões sazonais de correlação entre o índice SST do Atlântico sul e as vazões no mês de março, com defasagem de um a dois meses. As mesmas informações são apresentadas para o mês de setembro, na Figura 6.8. Os círculos

em azul (vermelho) apresentam os resultados em que as correlações são negativas (positivas).

Durante a estação chuvosa (março), uma estrutura em formato de "gangorra" é observada, em que os reservatórios localizados entre as latitudes 0º e 22°S apresentam séries de vazão negativamente correlacionadas com o índice SST, enquanto que os reservatórios que se encontram entre as latitudes 22°S e 33°S são correlacionados positivamente. Essa estrutura é semelhante à verificada em Lima e Lall (2010).

Entende-se que a forte (fraca) convecção na ZCAS resulta em abundante (escassa) precipitação em uma faixa que se estende desde a Amazônia ao sul do Atlântico e, também, em chuvas escassas (abundantes) na Argentina, Chile e sul do Brasil (Lima e Lall, 2008-b e 2009). Anomalias negativas na SST do Atlântico sul podem ser responsáveis por aumentar a alta do Atlântico subtropical sul e, consequentemente, intensificar a ZCAS. Por outro lado, anomalias positivas da SST do Atlântico sul podem diminuir ainda mais a já fraca estação seca na ZCAS e favorecer as elevadas vazões na maioria dos reservatórios (Lima e Lall, 2010).

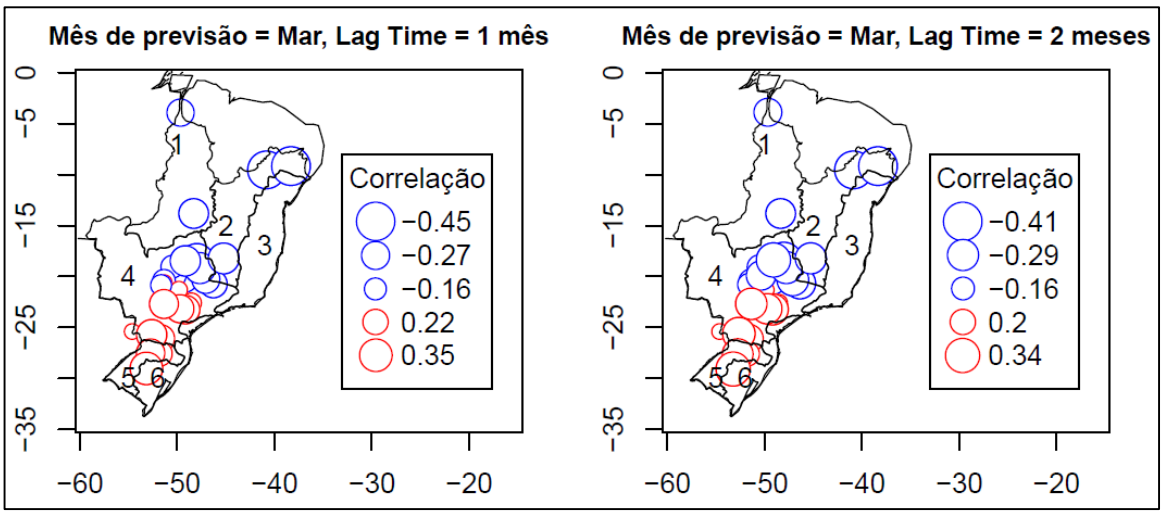

Figura 6.7 - Correlações entre as vazões do mês de março e o índice SST, com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

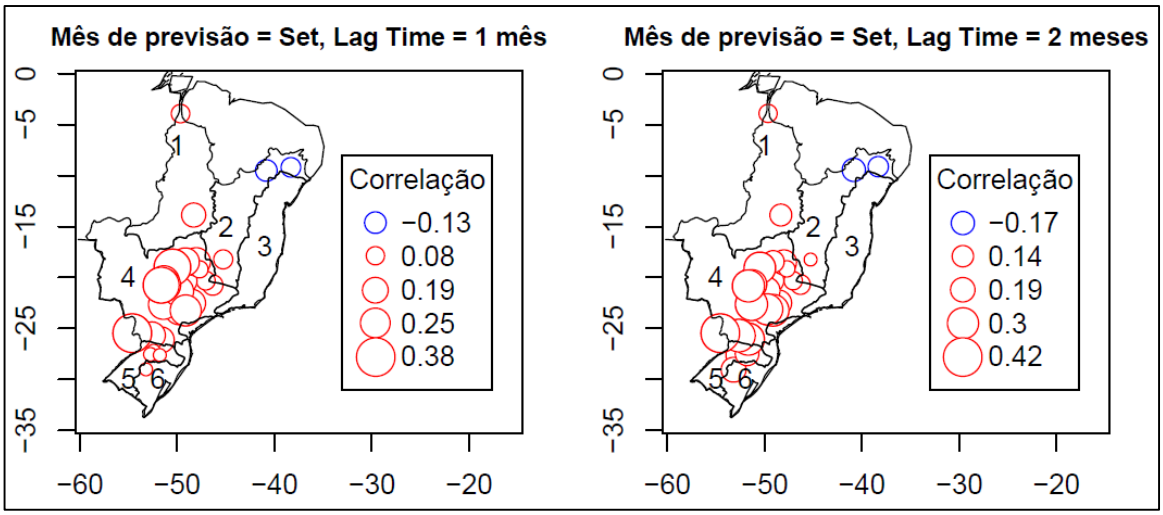

Figura 6.8 - Correlações entre as vazões do mês de setembro e o índice SST, com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

A Figura 6.9 apresenta os resultados médios anuais das correlações entre as vazões e o preditor SST, onde-se é mantido o formato de gangorra, com correlações positivas para os reservatórios que se encontram entre as latitudes 0 e 22^0S e negativas, entre as latitudes 22^0S e 33^0S.

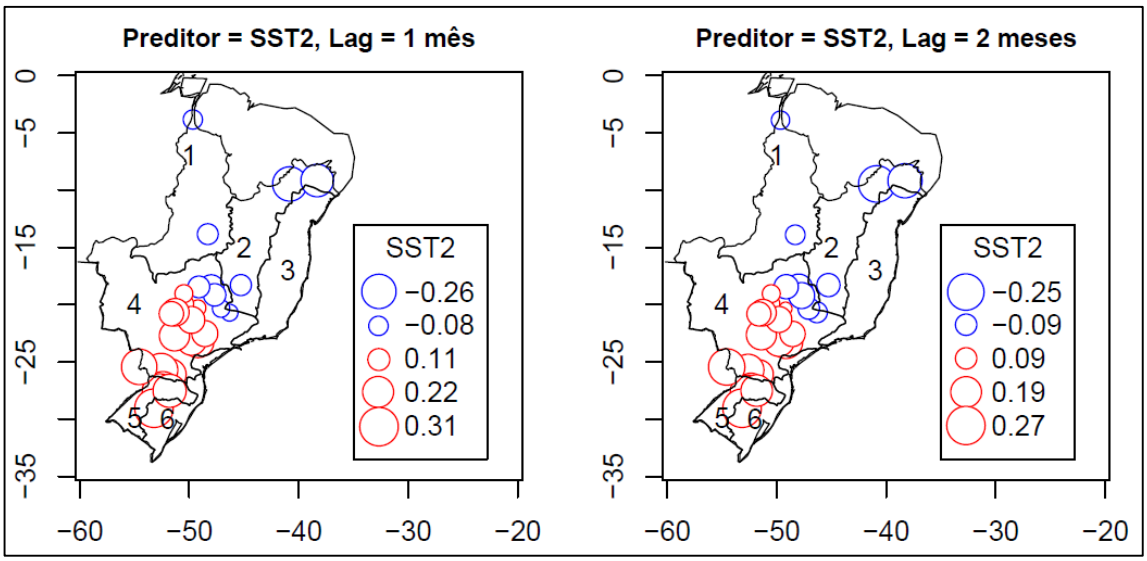

Figura 6.9 - Valores médios anuais das correlações entre as vazões e o índice SST, com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

As correlações médias anuais entre as vazões e o índice do vento zonal com um a dois meses de defasagem são exibidas na Figura 6.10, apresentando comportamento aproximadamente inverso ao do índice SST, pois aqui as séries de vazão apresentam correlações positivas entre as latitudes 0º e 22°S e negativas entre as latitudes 22°S e 33°S.

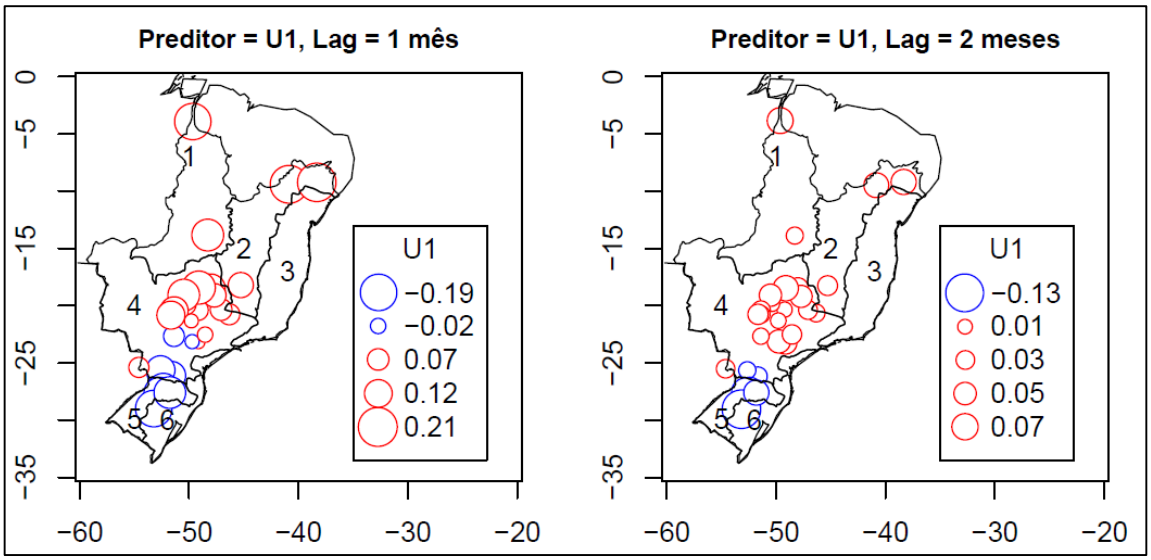

Figura 6.10 - Valores médios anuais das correlações entre as vazões e o índice de vento zonal (U1), com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

Por fim, os diagramas das funções de autocorrelação dos vinte e oito reservatórios apresentados na Figura 6.11 mostram que a persistência é um termo importante para modelar o processo de vazões, em particular durante o período seco. Em um painel não apresentado aqui, conclui-se que no período chuvoso, a persistência é inferior a dois meses nos reservatórios do nordeste e centro-oeste. No período seco (setembro), a persistência é superior a dois meses em todo o território.

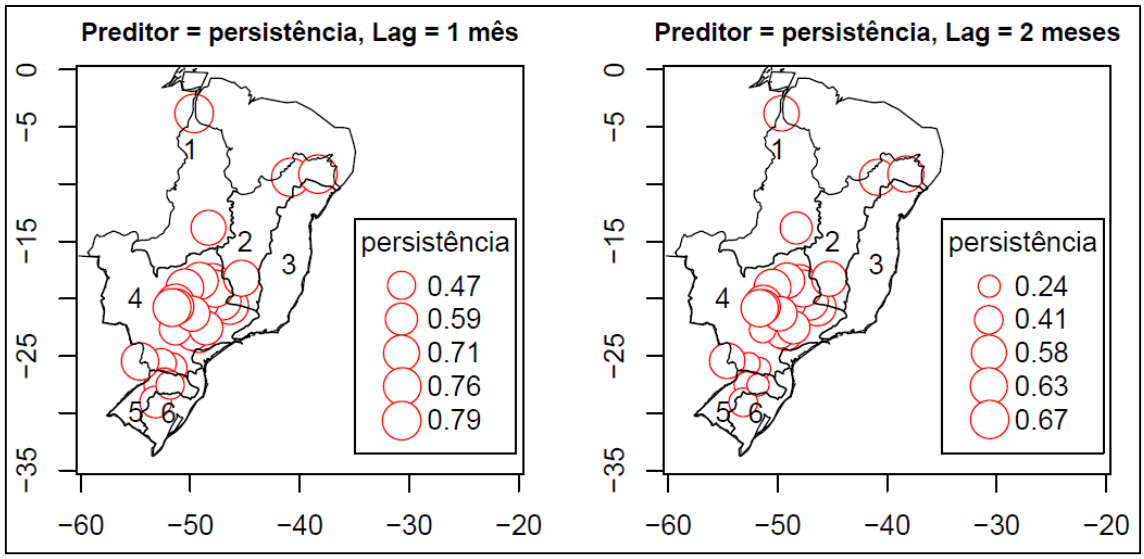

Figura 6.11 - Funções periódicas de auto-correlação com *lags* de um e dois meses. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

O critério utilizado nesta dissertação para avaliar e comparar a qualidade do modelo mensal foi o coeficiente de Nash-Sutcliffe (NS), que permite medir as mudanças no desempenho de previsão como um acréscimo na persistência e nos preditores meteorológicos em um modelo de média periódica. Note-se que NS pode variar entre $-\infty$ e +1, e sempre que $0 < NS \leq 1$ o modelo testado tem desempenho melhor do que o modelo mais simples possível (o modelo de média de longo termo). Quando NS<0 entende-se ser preferível utilizar a média de longo termo das observações do que os resultados previstos pelo modelo.

As Figuras 6.12 e 6.13 apresentam a distribuição espacial dos índices NS do modelo PARX para os meses de março e setembro, respectivamente. No período chuvoso (março), verificam-se aglomerados de reservatórios de baixa previsibilidade na região de transição entre o centro-leste e sul do Brasil (entre 18ºS e 25ºS de latitude). No período seco (setembro), essas regiões apresentam maior previsibilidade. A vazão de setembro mantém sua previsibilidade durante os dois *lags*, para os reservatórios no nordeste e parte do centro-leste do Brasil. A previsibilidade em março para a vazão de reservatórios localizados na região entre 20º e 25º de latitude irá aumentar somente com atrasos (*lags*) de três meses (comparativo não apresentado nesta dissertação).

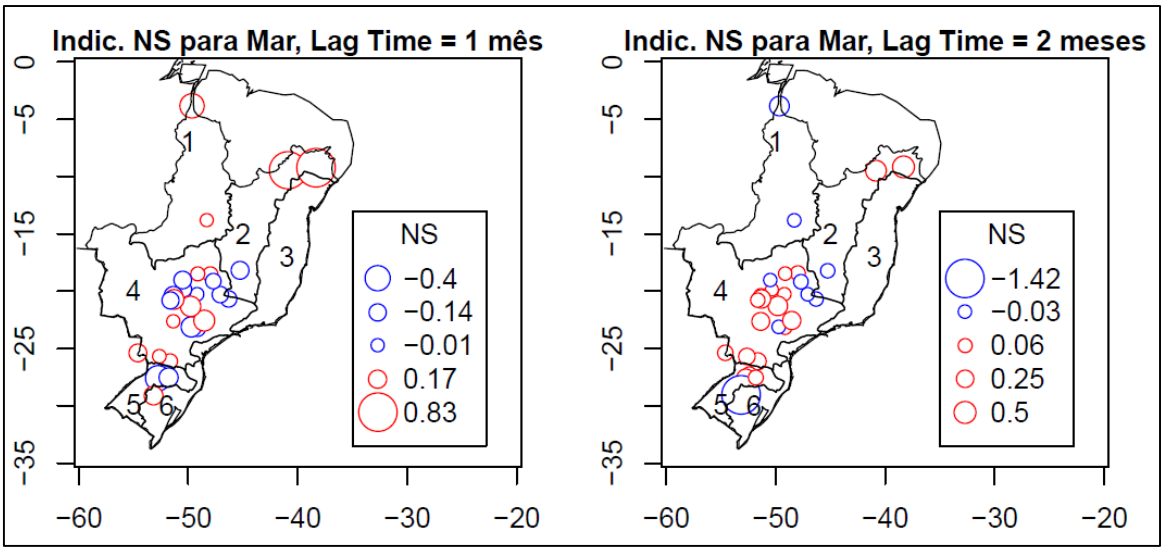

Figura 6.12 - Distribuição espacial dos índices NS para o modelo PARX com *lags* de um a dois meses, no mês de março. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

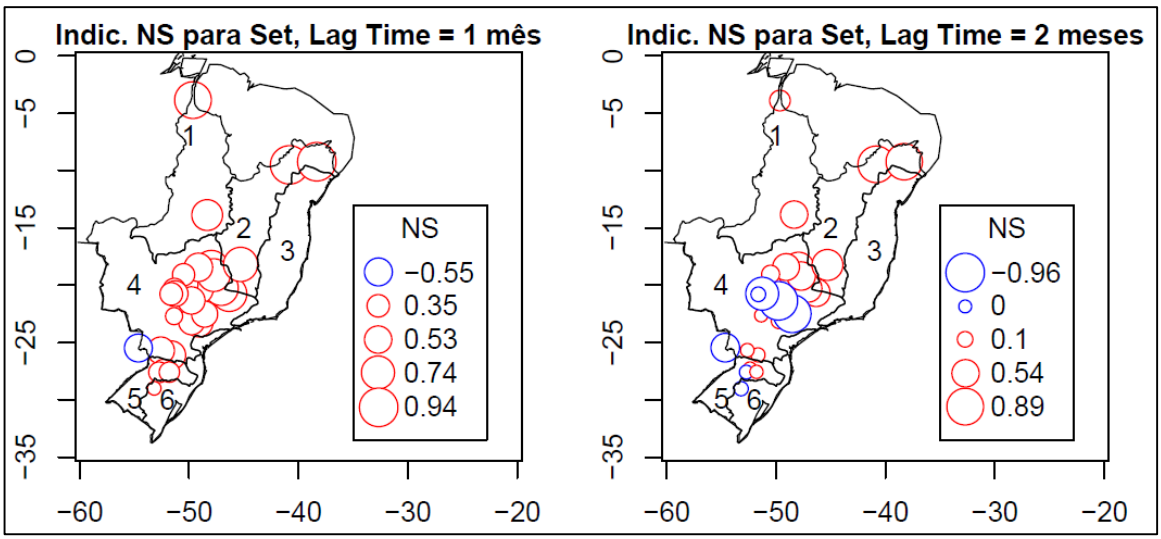

Figura 6.13 - Assim como na Figura 6.12, mas para o mês de setembro. Fonte: elaboração própria, adaptado de Lima e Lall (2010).

A Figura 6.14 apresenta um comparativo entre os resultados globais do Indicador NS para os modelos mensais PARX e PAR com *lags* de um mês e a Figura 6.15 traz os resultados quando o *lag* mensal é igual a dois. Embora haja uma natural redução do desempenho do modelo à medida que o *lag* mensal aumenta, verifica-se que o indicador global NS do modelo PARX, que utiliza indicadores climáticos como preditores exógenos, é superior aos resultados do modelo PAR, sobretudo com *lag* de 1 mês.

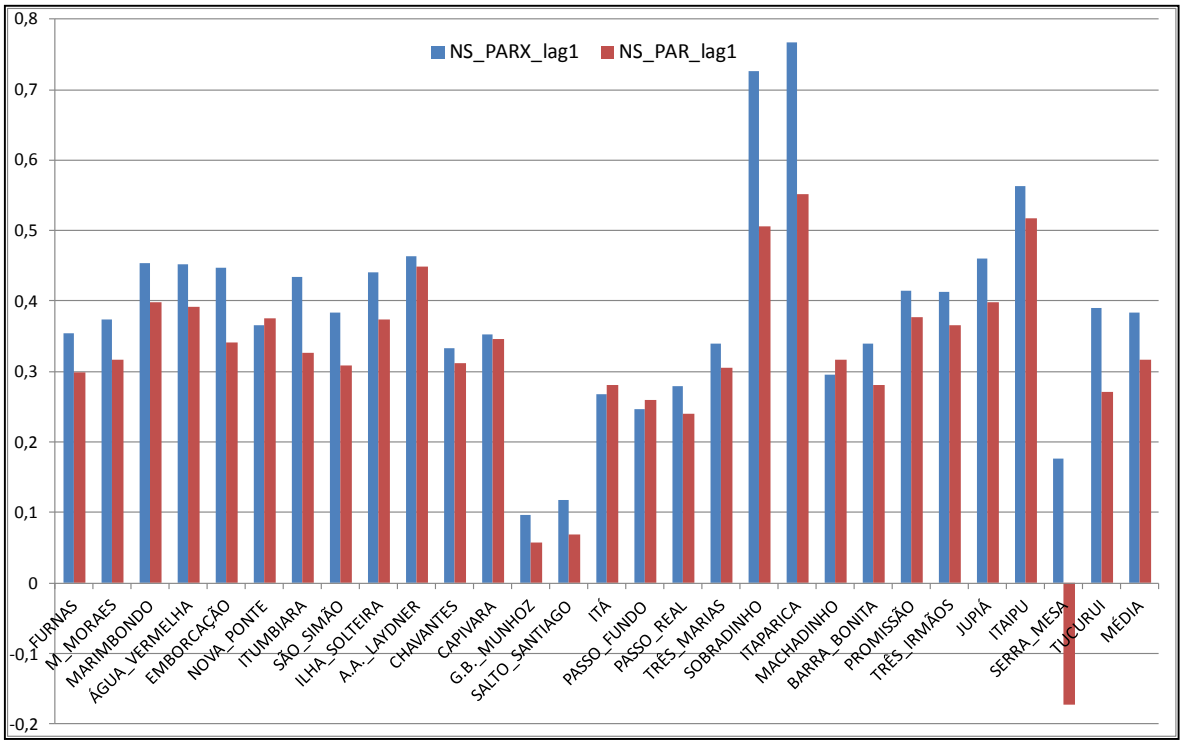

Figura 6.14 - Comparativo entre os indicadores globais NS dos modelos PARX e PAR com *lag* de um mês.

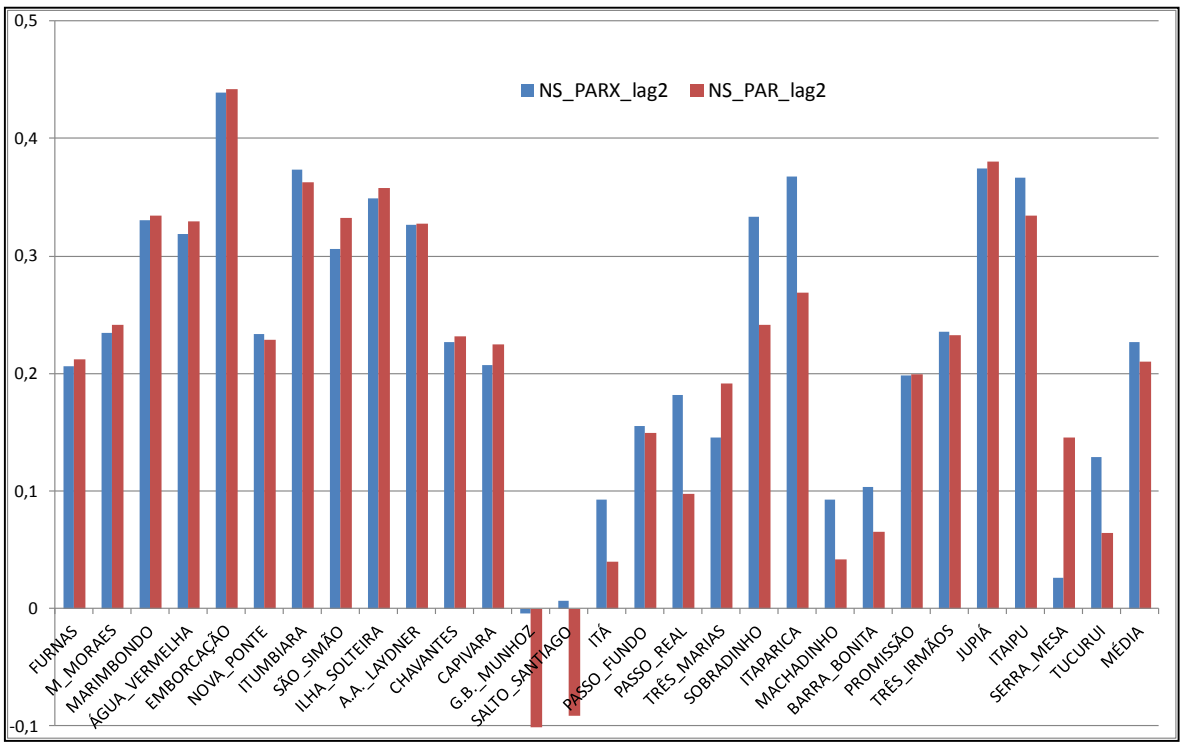

Figura 6.15 - Assim como na Figura 6.14, mas para o *lag* de dois meses.

6.2. RESULTADOS PARA AS MODELAGENS SEMANAIS - DESEMPENHO DOS MODELOS DO ONS.

Os resultados dos indicadores das previsões de vazão para as usinas de diversas bacias hidrográficas brasileiras com uma a três semanas de antecedência são disponibilizados no Relatório Anual de Avaliação das Previsões de Vazões, emitidos regularmente pelo ONS desde 2007, mas com informações detalhadas a partir de 2009. As Figuras 6.16 a 6.18 ilustram os resultados médios obtidos entre os anos de 2009 e 2013 pelos modelos utilizados pelo ONS nas 28 usinas estudadas para o indicador DM. A partir de 2011, o ONS passou a disponibilizar o resultado dos indicadores discretizados trimestralmente, permitindo atestar o desempenho dos modelos em períodos hidrológicos distintos. Os Relatórios Anuais de Avaliação das Previsões de Vazões estão disponíveis em www.ons.org.br, sendo que os dados utilizados para confecção das Figuras 6.16 a 6.18 constam do Apêndice C desta dissertação.

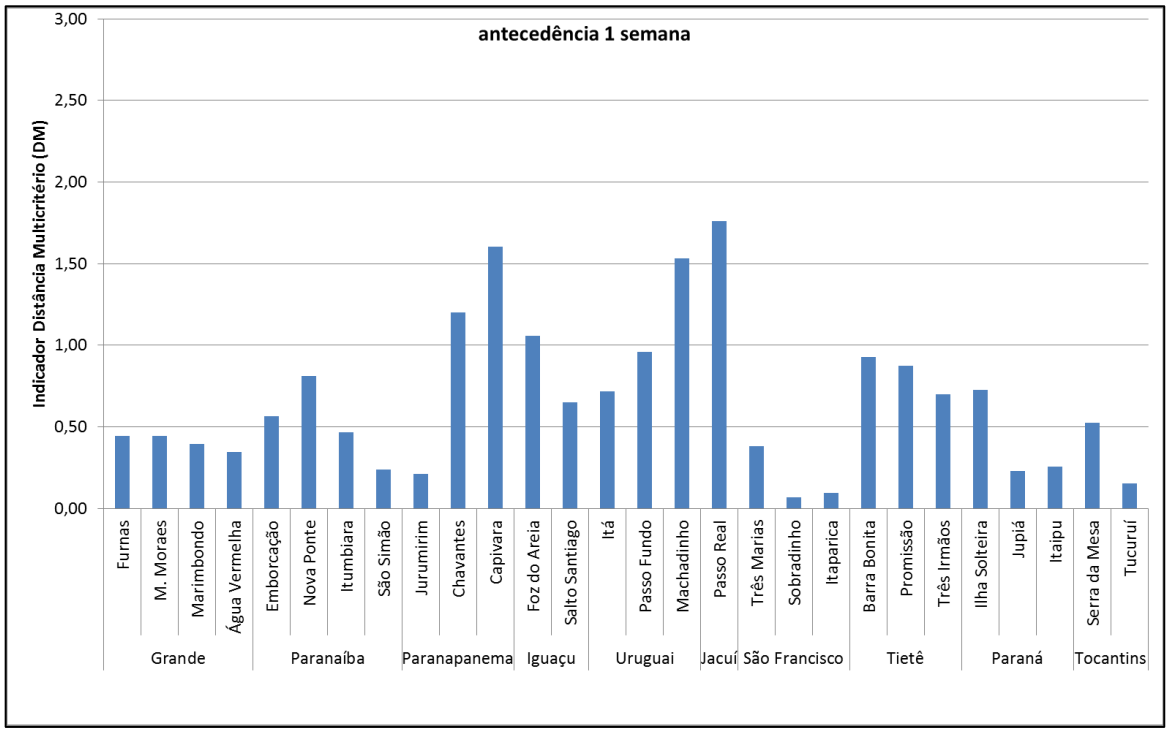

Figura 6.16 - Resultados do indicador distância multicritério para os modelos utilizados pelo ONS com antecedência de 1 semana (média do período entre 2009 e 2013).

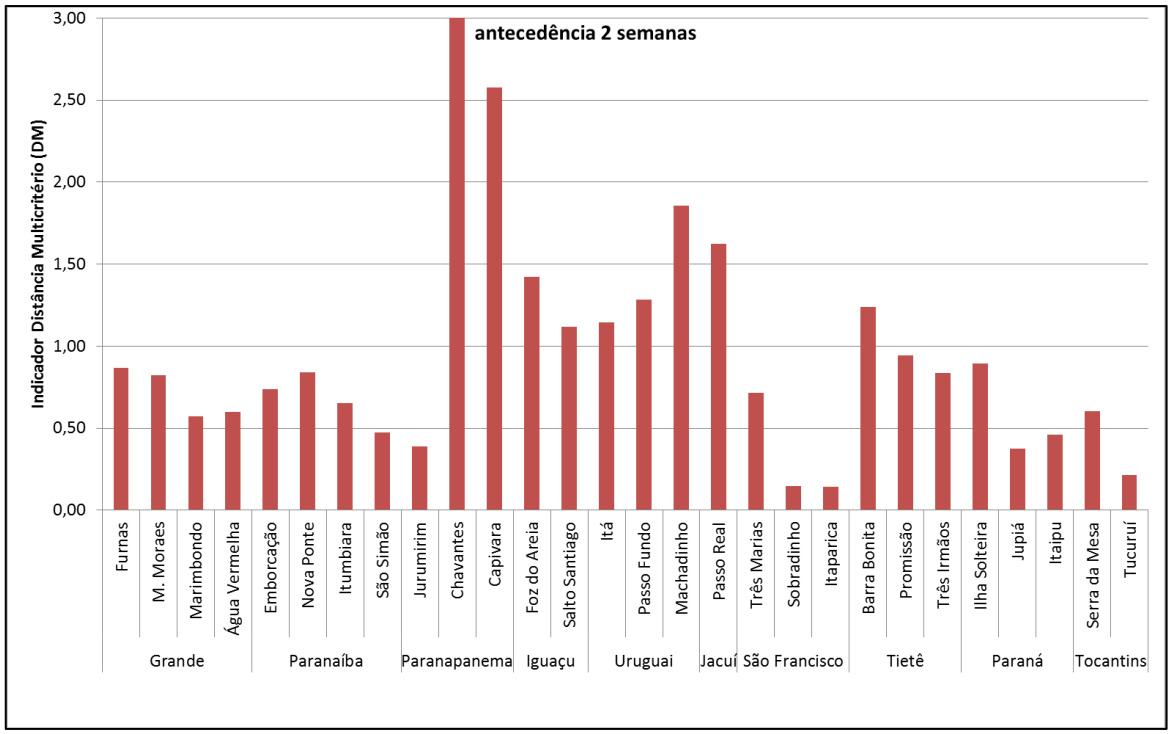

Figura 6.17 – Assim como na Figura 6.16, mas para a antecedências de 2 semanas.

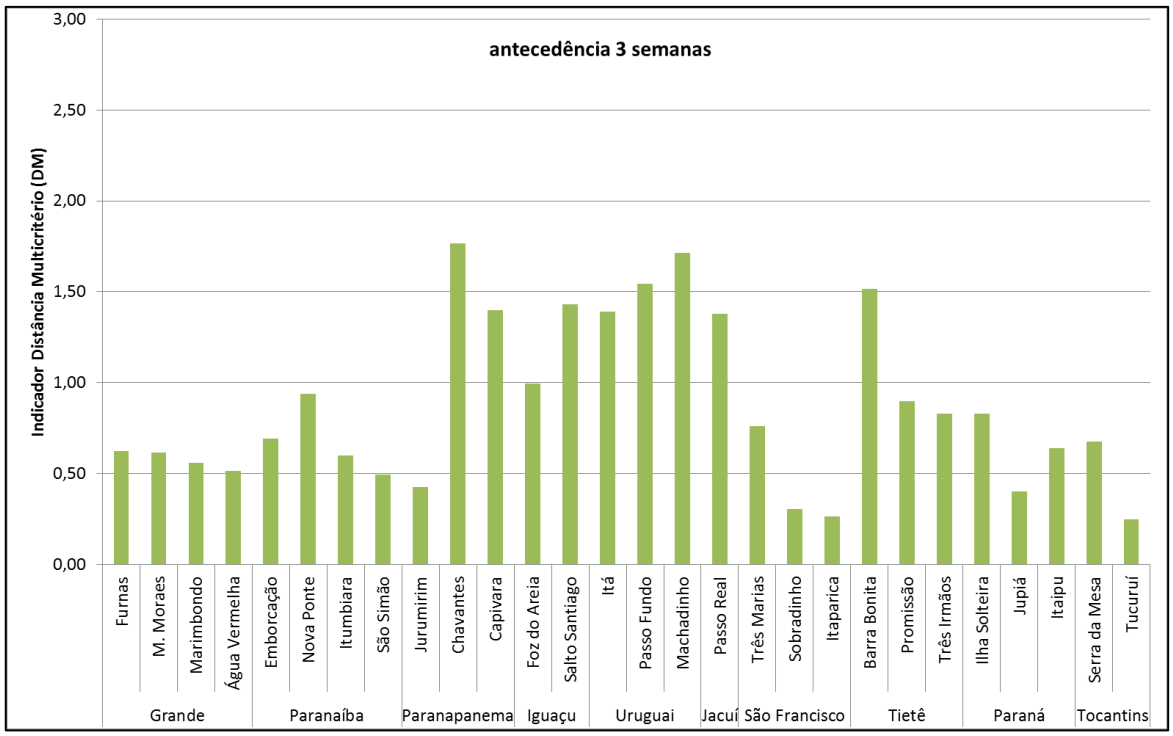

Figura 6.18 - Assim como na Figura 6.16, mas para a antecedências de 3 semanas.

Ressalta-se que, diferente dos resultados apresentados nos Relatórios Anuais de Previsão de Vazões do ONS, que são calculados ano a ano (vide Apêndice C), os

indicadores calculados nesta dissertação utilizaram, tanto para os resultados do ONS quanto para os obtidos nas ponderações, os valores de todo o período estudado, ou seja, de janeiro de 2009 a setembro de 2014.

6.3. RESULTADOS PARA AS MODELAGENS SEMANAIS - PONDERAÇÃO BAYESIANA DOS MODELOS.

Após a etapa de modelagem mensal e de aquisição dos dados oriundos dos modelos semanais do ONS, os dados resultantes dos modelos foram submetidos à técnica de ponderação bayesiana de modelos (BMA).

Para cada ano disponibilizado pelo ONS, de 2009 a 2014, procedeu-se a uma validação cruzada, ou seja, as amostras foram divididas em dois períodos: um período de treinamento, para fins de estimação dos parâmetros do modelo, e um período de validação do modelo calibrado. Assim, a ponderação BMA para os valores previstos no ano de 2009, por exemplo, não contam com o referido ano no período de treinamento (calibração), que foi composto pelos demais anos (2010 a 2014). O mesmo procedimento foi estabelecido para todos os anos do estudo. Com isso, o número de amostras para os períodos de treinamento e de validação ficou distribuído de acordo com a Tabela 6.3, em função das semanas de antecedência nas previsões.

Tabela 6.3 - Número de amostras no período de treinamento e de validação.

Semana(s) de antecedência	**Amostras (Período de treinamento)**	**Amostras (Período de validação)**	**Total de amostras (treinamento + validação)**
Uma	249	52	301
Duas	249	52	301
Três	228	48	276
Quatro	171	36	207
Cinco	114	24	138
Seis	57	12	69

O número de amostras apresentado na Tabela 6.3 é válido para os anos completos (de 2009 a 2013), já que no ano de 2014 o período analisado compreendeu apenas os meses de janeiro a setembro. Assim, 2014 possui um menor número de amostras no

período de validação que os demais e, consequentemente, um maior número de amostras no período de treinamento.

6.3.1. Resultado das previsões dos vinte e oito reservatórios em estudo.

Conforme já mencionado, o procedimento BMA foi aplicado para os vinte e oito principais reservatórios do sistema elétrico brasileiro. Na Figura 6.19, apresentam-se os resultados iniciais da previsão de uma a seis semanas de antecedência para a UHE Emborcação, uma das usinas analisadas, escolhida para fins de ilustração da metodologia adotada. Nota-se que, com o aumento da semana de antecedência, o modelo semanal vai perdendo importância em detrimento do modelo mensal. Na transição da terceira para a quarta semana o peso do modelo mensal supera o peso do modelo semanal.

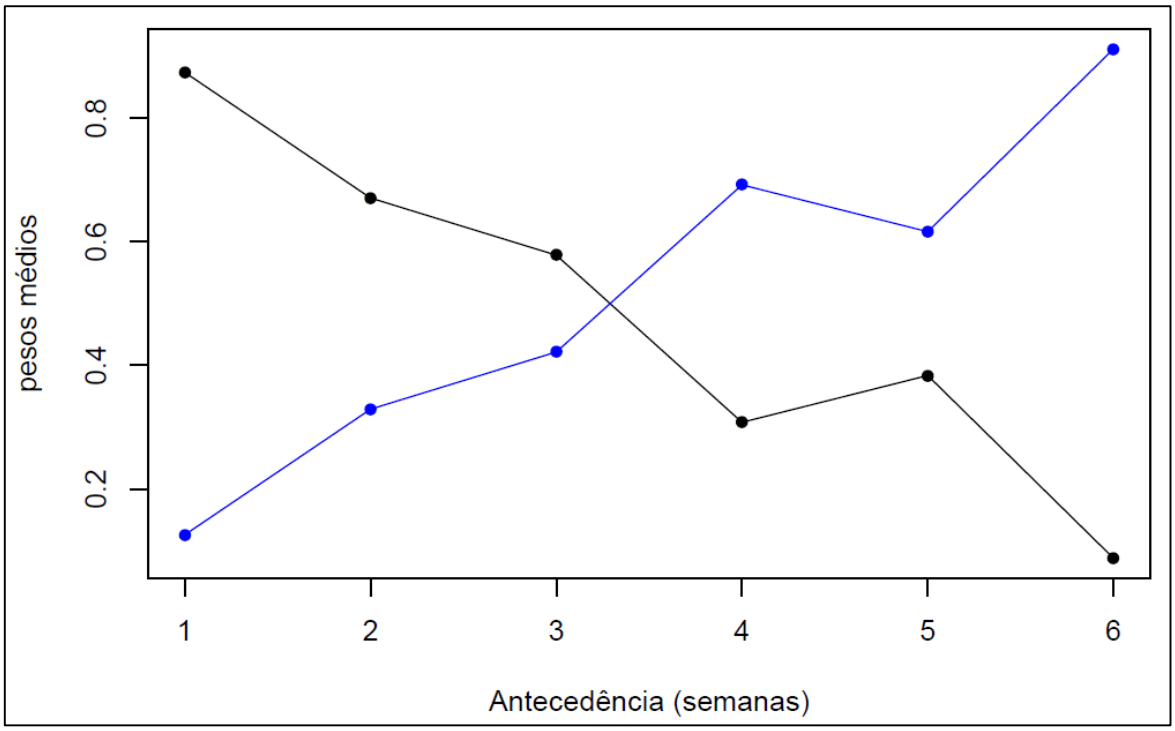

Figura 6.19 - Distribuição dos pesos Bayesianos para a UHE Emborcação para cada antecedência semanal (período completo). A linha na cor azul refere-se aos pesos mensais e a linha na cor preta, aos semanais.

Da quarta para a quinta semana, o modelo semanal recupera parte de sua importância nas ponderações, pois o modelo mensal utilizado passa a ser o modelo de previsão com dois meses de antecedência, que possui desempenho inferior às previsões realizadas com um mês de antecedência, retomando o seu crescimento da quinta para a sexta semana. A Figura 6.19 ilustra a transição dos pesos para o período completo analisado (janeiro de 2009 a setembro de 2014), deixando claro o momento em que o modelo mensal passa a ter maior importância nas ponderações, lembrando que a soma dos pesos mensal e semanal sempre se iguala a um.

De forma a avaliar o comportamento da mistura de modelos de previsão para os reservatórios do estudo, a distribuição dos pesos obtidos para os modelos mensais em todos os reservatórios foi plotada espacialmente para os diversos horizontes de previsão. Os círculos em verde representam os pesos dos modelos mensais que não superaram os pesos semanais (menores que 0,5); os círculos na cor vermelha indicam o contrário (pesos mensais maiores que 0,5). A Figura 6.20 apresenta os resultados para as primeiras duas semanas de antecedência. Em virtude da persistência das vazões, os modelos semanais sobrepõem-se aos mensais para a antecedência de uma semana em todos os reservatórios testados. Com antecedência de duas semanas, os modelos semanais ainda são soberanos nas ponderações, exceção feita para as previsões de afluências aos reservatórios A.A. Laydner, Passo Fundo e Passo Real.

A Figura 6.21 ilustra a distribuição espacial dos pesos mensais para a terceira e quarta semanas de antecedência. Nota-se um equilíbrio entre os reservatórios que utilizam os modelos semanais e mensais, sobretudo quando a antecedência atinge quatro semanas. Os modelos semanais são preponderantes nas usinas localizadas acima da latitude 20ºS e os modelos mensais, nas usinas abaixo dessa linha divisória, indicando que as usinas abaixo da latitude 20ºS possuem melhor resposta às previsões com o modelo mensal (ou seja, possuem um sinal climático mais influente nas previsões).

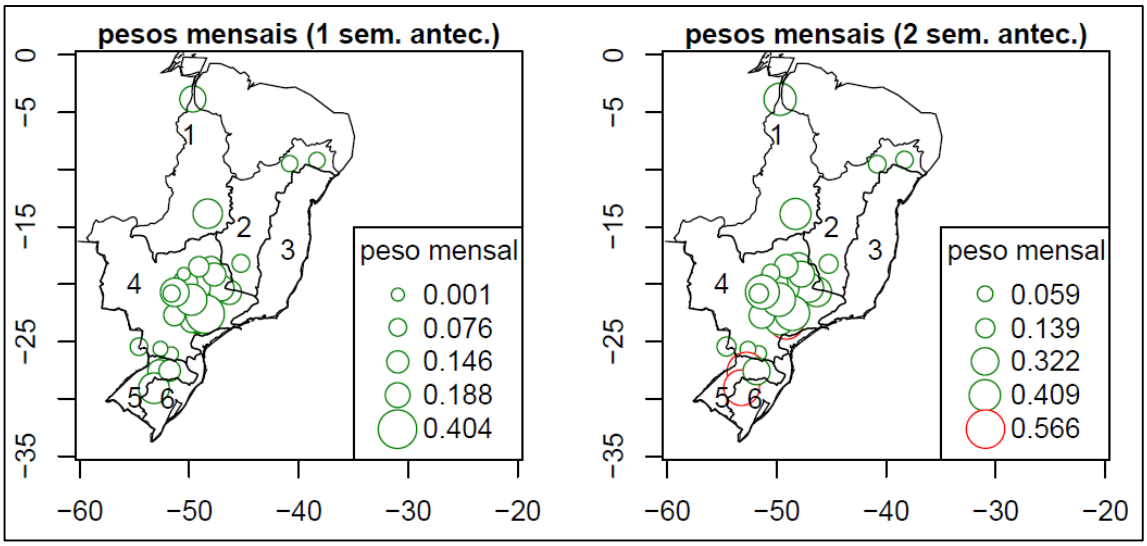

Figura 6.20 - Distribuição espacial dos pesos dos modelos de previsão mensal para os 28 reservatórios analisados, nas previsões com antecedências de uma (esquerda) e duas (direita) semanas. Valores na cor verde indicam que o peso dos modelos semanais é maior que os pesos dos modelos mensais, os valores na cor vermelha, indicam o contrário.

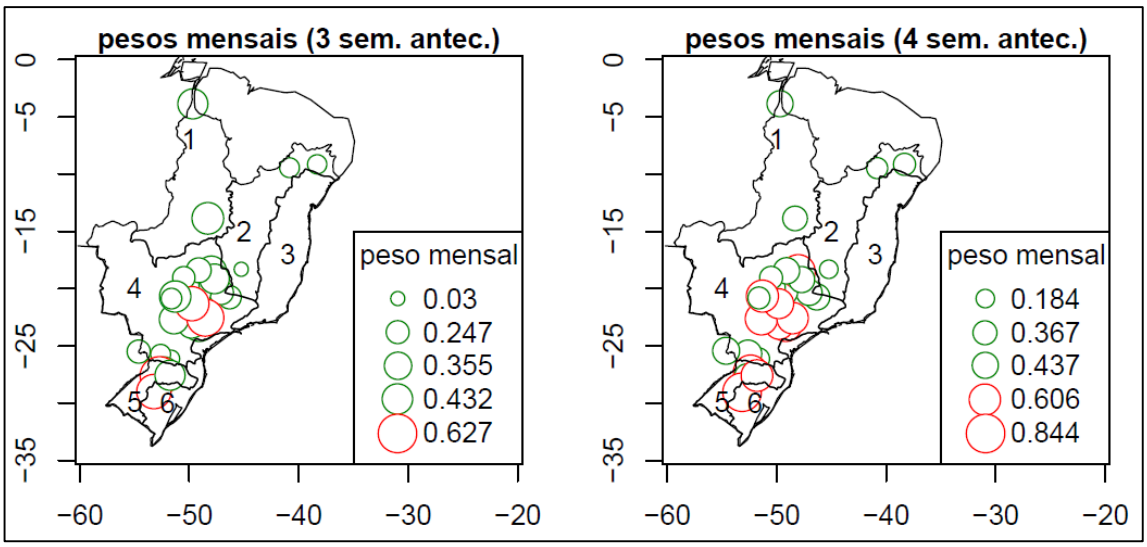

Figura 6.21 – Assim como na Figura 6.20, mas para previsões realizadas com três (esquerda) e quatro (direita) semanas de antecedência.

Ainda se utilizando das previsões da UHE Emborcação como exemplo, verificou-se que os indicadores mensais da aplicação do modelo PARX mensal, além de manterem sua qualidade no decorrer das semanas, obtiveram melhoria na maioria dos casos, o que pode ter contribuído para sua grande utilização nas ponderações BMA semanais.

A Figura 6.22 traz o resultado dos indicadores mensais de desempenho do modelo PARX para a UHE Emborcação no período entre 2009 e 2014, onde são comparadas as previsões mensais (convertidas para a escala semanal) e os valores semanais observados. Verifica-se, para a referida usina, uma melhoria nos quatro indicadores (redução no RMSE, MAPE e DM) e acréscimo no índice NS para o modelo mensal. Com isso, o indicador DM passou de 0,65 na primeira semana para 0,45 na última. O ganho de qualidade de previsão no decorrer das semanas pode ter influenciando no aumento do peso bayesiano mensal nas ponderações.

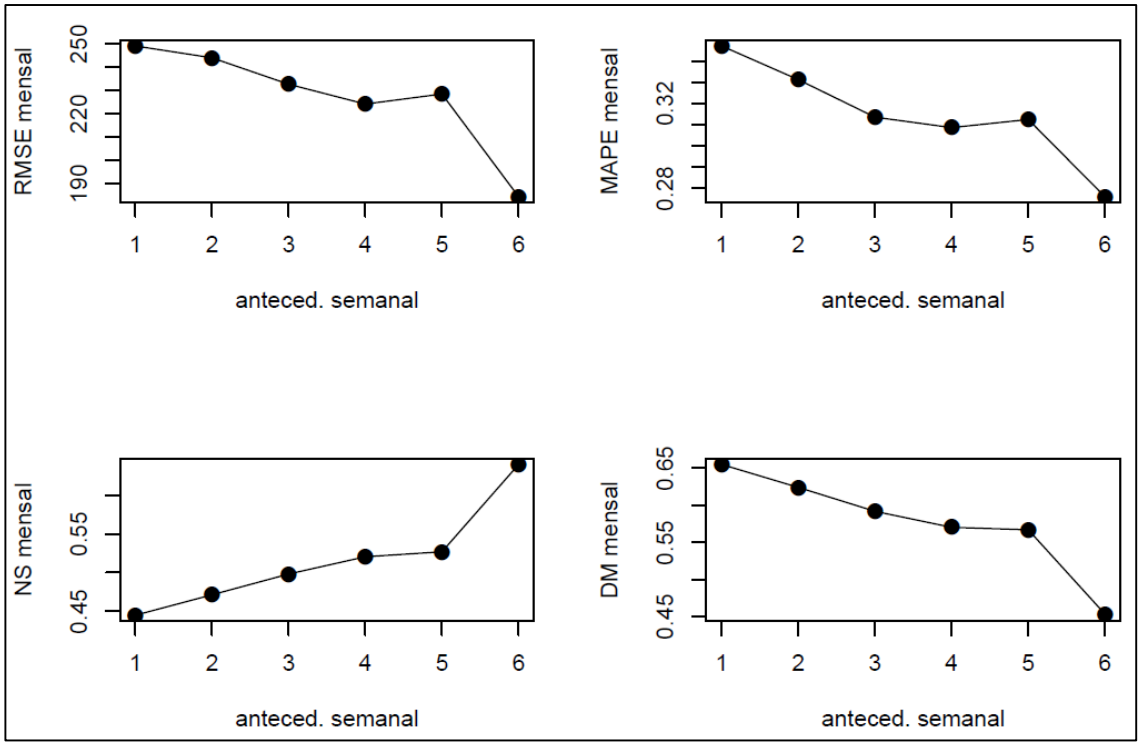

Figura 6.22 - Indicadores de desempenho para previsões realizadas com o modelo mensal de uma a seis semanas de antecedência para as afluências à UHE Emborcação.

O indicador distância multicritério (DM) de qualidade das previsões semanais no período completo estudado, traçado na Figura 6.23, mostra a melhoria das previsões semanais na UHE Emborcação em todas as semanas de antecedência, após a aplicação da técnica BMA, com melhores resultados na terceira, quarta e sexta semanas. O painel da esquerda mostra os resultados para o indicador DM. O painel da direita mostra as diferenças entre os resultados do indicador DM em cada semana (curva em azul – resultados dos modelos adotados pelo ONS; curva em vermelho – resultados após as

poderações). Os valores positivos mostram as semanas em que os indicadores DM das previsões semanais realizadas pelo ONS foram maiores que os das previsões BMA.

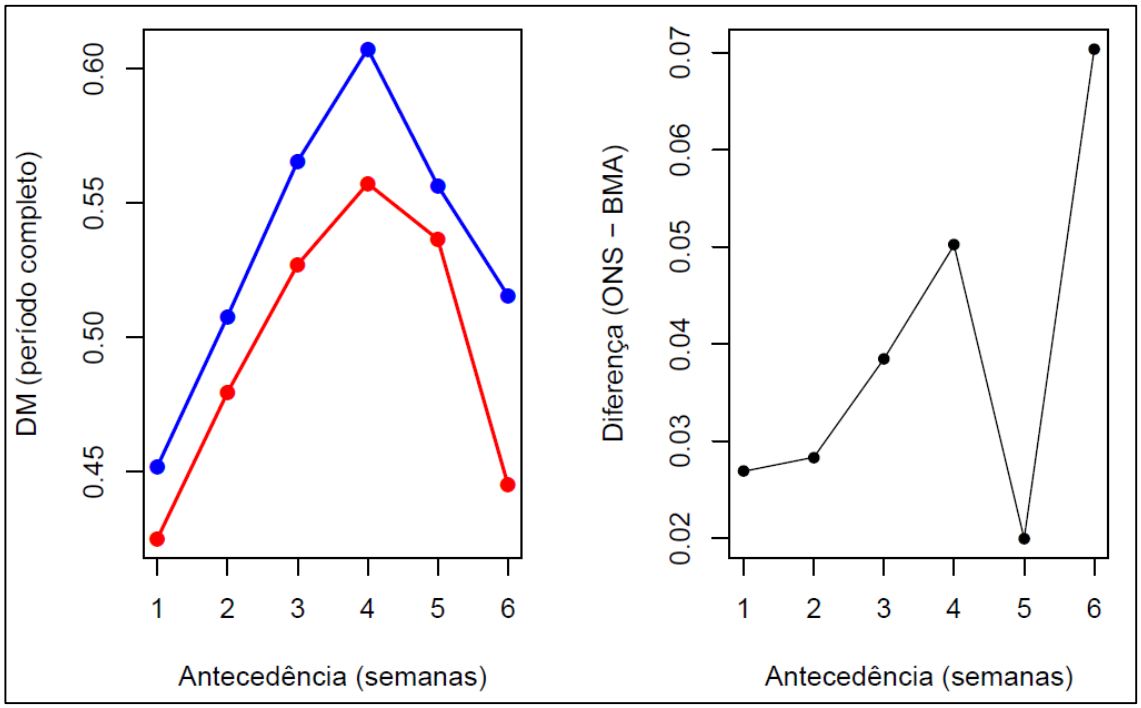

Figura 6.23 - Valores para o indicador DM com uma a seis semanas de antecedência na UHE Emborcação: o painel da esquerda mostra os resultados deste indicador para as previsões semanais realizadas pelo ONS (curva na cor azul) e das previsões BMA (curva vermelha) no período completo (entre 2009 e 2014) DM no período; o painel da direita mostra as diferenças entre os indicadores DM das previsões semanais para cada semana, antes e após a aplicação BMA.

A Figura 6.24 apresenta a distribuição dos pesos mensais para a quinta e sexta semana de antecedência. Nota-se, em alguns casos, a recuperação dos modelos semanais na ponderação bayesiana dos modelos. Também é possível verificar que os reservatórios de Sobradinho e Itaparica, na bacia do São Francisco, passam a utilizar o modelo mensal sobremaneira na sexta semana.

Cabe destacar que alguns reservatórios possuem redução nos pesos do modelo mensal quando a antecedência das previsões atinge cinco e seis semanas. Com isso, o modelo PARX mensal contribui com menor intensidade para as ponderações BMA. A Figura 6.25 apresenta, a título ilustrativo, a tendência dos indicadores mensais com antecedências de uma a seis semanas para a UHE Tucuruí. Nota-se a redução do

desempenho do modelo mensal com dois meses de antecedência (quinta e sexta semanas), conforme figura a seguir, que compara as previsões mensais (convertida para a escala semanal) e os valores semanais verificados.

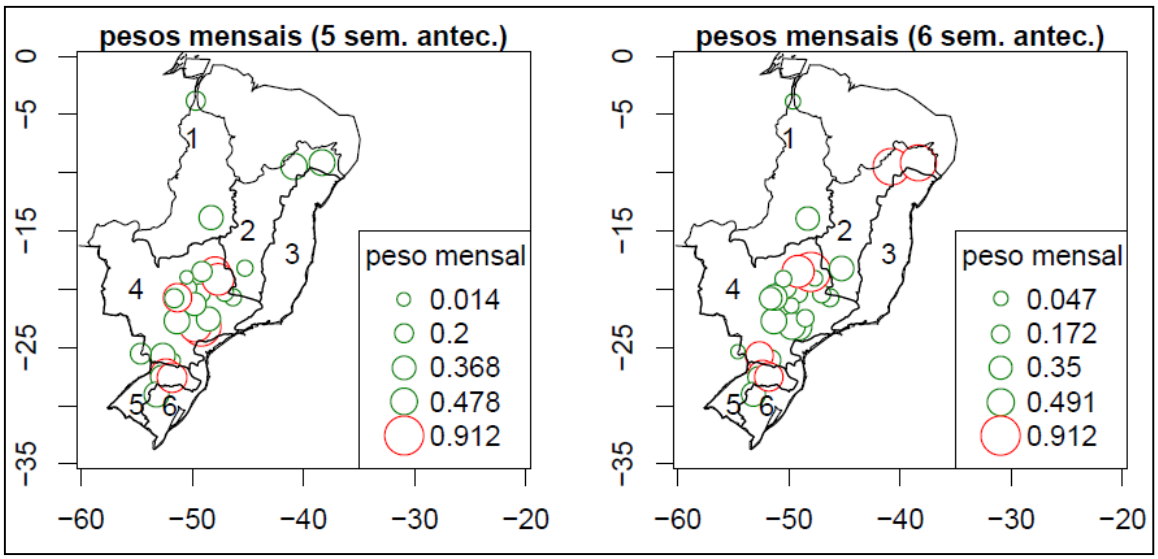

Figura 6.24 – Assim como na Figura 6.20, mas para previsões realizadas com cinco (esquerda) e seis (direita) semanas de antecedência.

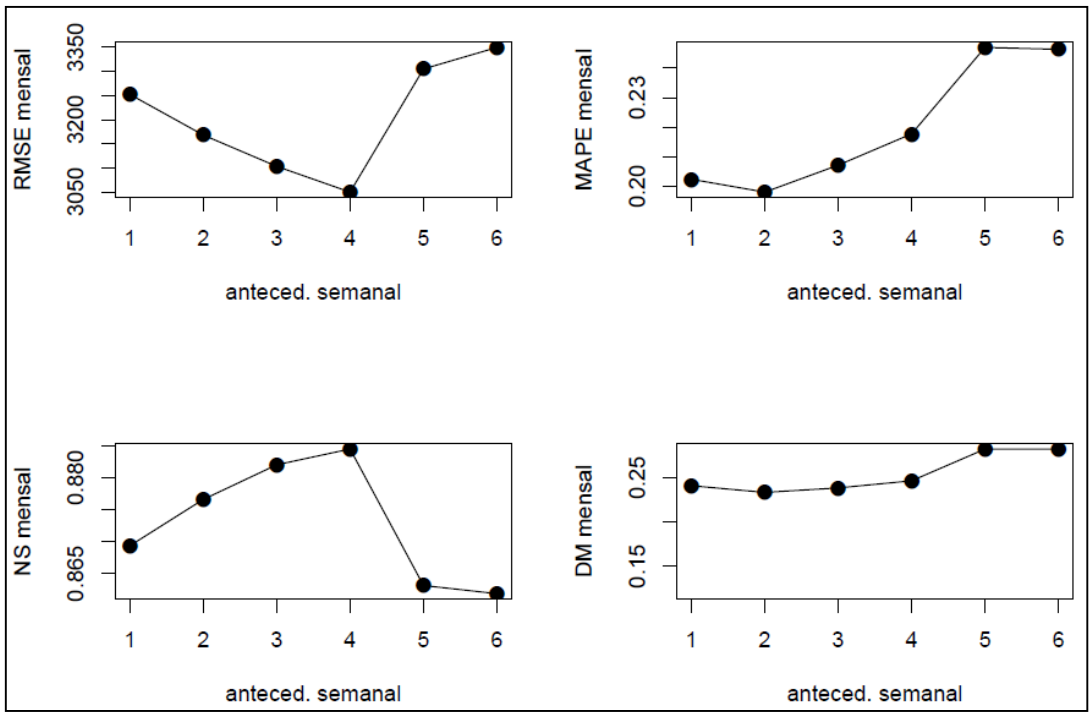

Figura 6.25 - Indicadores com uma a seis semanas de antecedência na UHE Tucuruí para o modelo mensal em relação aos valores verificados semanais (período completo).

Para o caso ilustrado (UHE Tucuruí), as perdas de desempenho do modelo mensal no segundo mês (quinta e sexta semanas) podem ter refletido na sua baixa utilização quando da aplicação da técnica BMA. Outra justificativa pode advir da grande redução do número de amostras para os modelos com cinco e seis semanas de antecedência. Amostras de menor tamanho possuem menos informação nos dados, com isso os valores das probabilidades a posteriori dos modelos (PMPs) – os pesos bayesianos – aproximam-se dos valores a priori, auferindo menores contribuições às ponderações que em modelos com maior número de amostras (Gibbons et al., 2008).

6.3.2. Resultados espacializados dos indicadores de qualidade das previsões

As Figuras 6.26 a 6.29 trazem os resultados espacializados para as diferenças percentuais entre os indicadores de qualidade das previsões semanais realizadas pelo ONS e as previsões semanais após a aplicação da técnica BMA. Valores em vermelho indicam que as diferenças entre as previsões semanais realizadas pelo ONS e as previsões BMA são positivas e valores em azul indicam o contrário. Os resultados dos indicadores encontram-se tabelados no Apêndice E desta dissertação.

Na Figura 6.26, que apresenta os resultados para o indicador RMSE, verifica-se que as diferenças percentuais entre as previsões semanais conduzidas pelo ONS e as previsões BMA são positivas para praticamente todos os reservatórios em todas as semanas de antecedência, ou seja, os erros das previsões semanais são maiores que os erros BMA. Nas duas últimas semanas, nota-se a redução no RMSE de alguns reservatórios. Dentre esses casos de redução, encontram-se algumas usinas nos rios Paranaíba e Grande.

A Figura 6.27 apresenta as diferenças percentuais para o indicador de curto prazo MAPE em todas as semanas de antecedência. Verifica-se um equilíbrio para este indicador em todas as antecedências, cujas diferenças estão em sua maioria entre -10% e +10%. O indicador MAPE, como informado, apresenta a vantagem de não ser influenciado pelo valor absoluto das vazões.

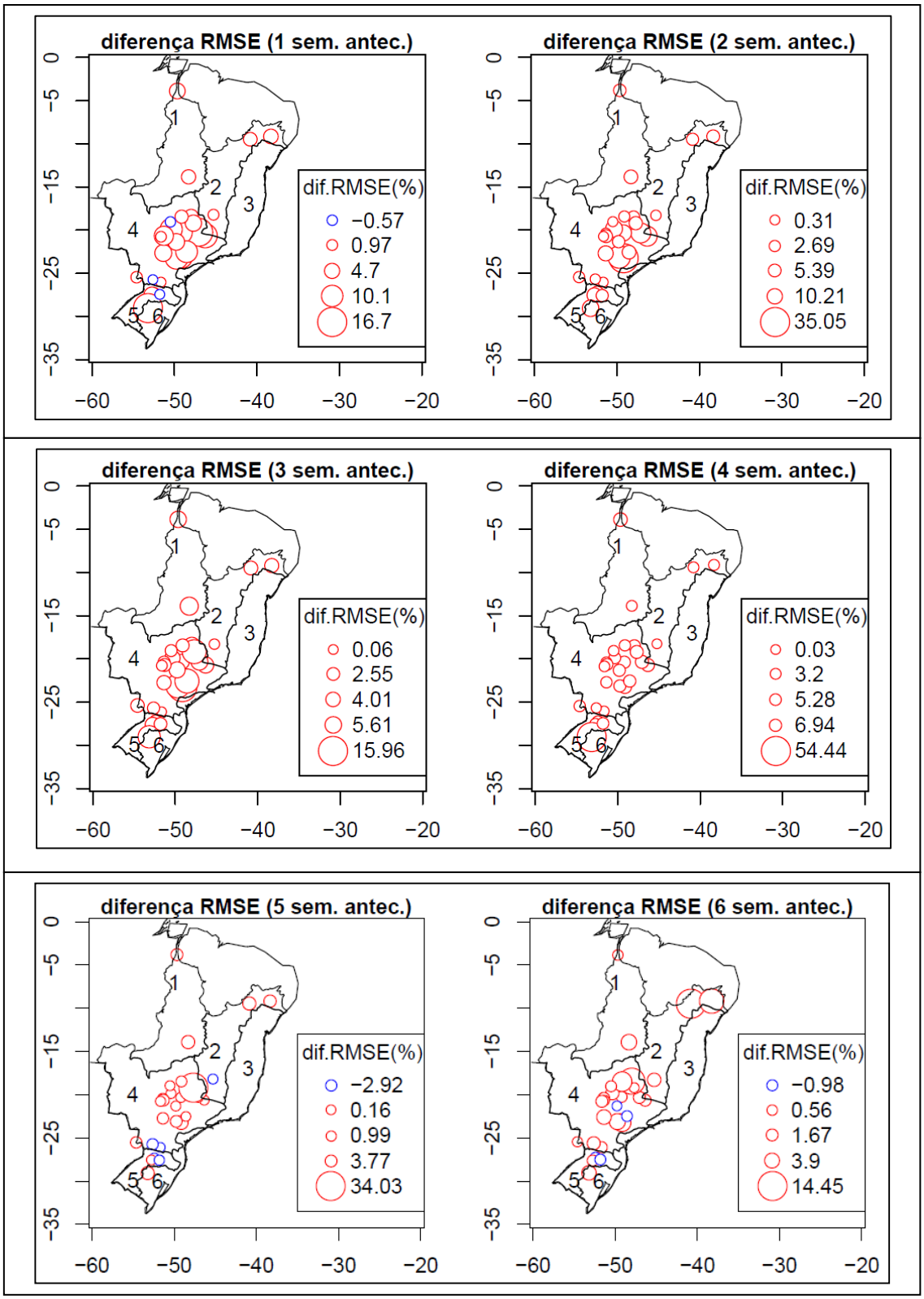

Figura 6.26 - Resultados das diferenças entre o indicador RMSE das previsões do modelo semanal e das previsões BMA para as seis semanas de antecedência.

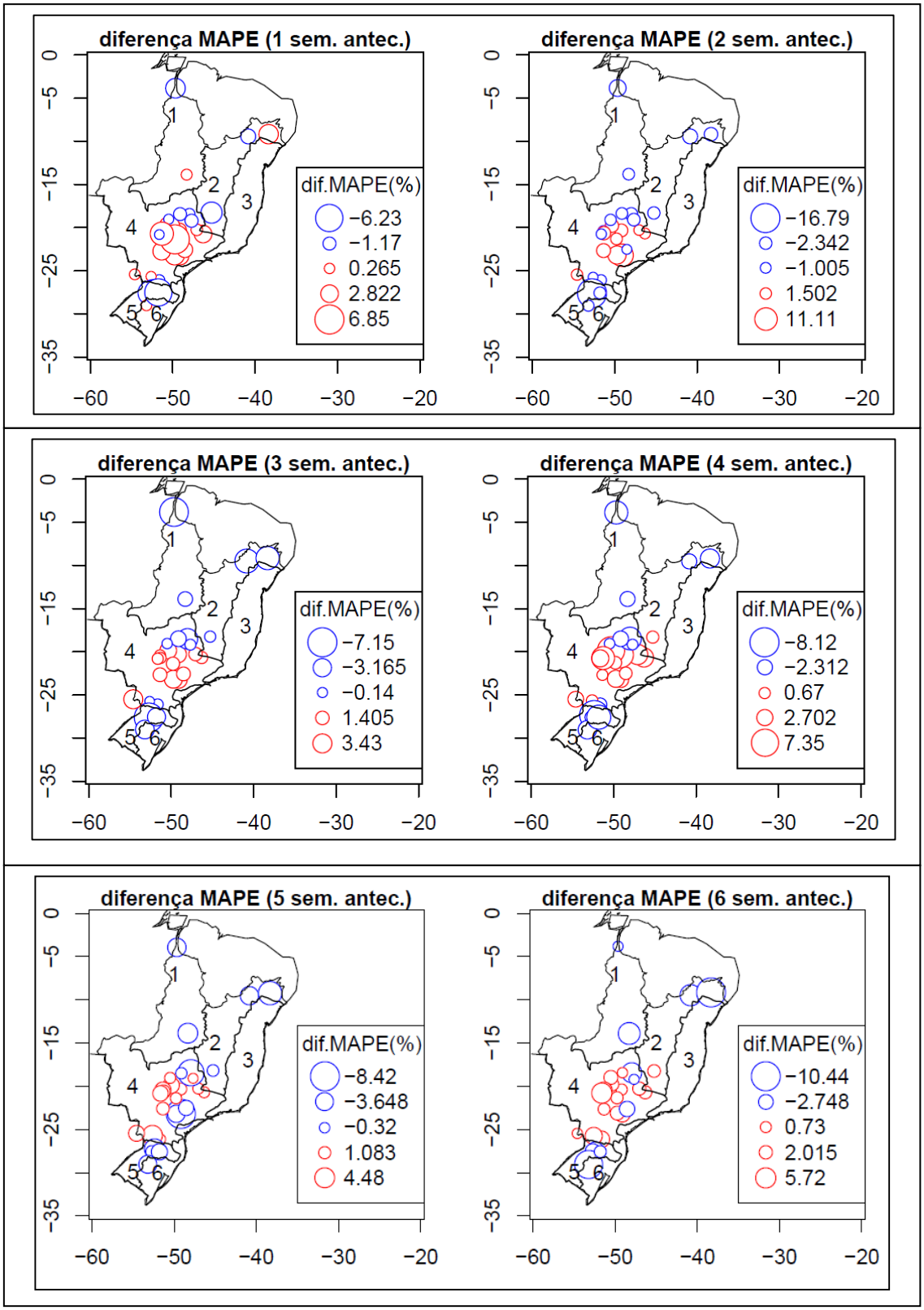

Figura 6.27 - Resultados das diferenças entre o indicador MAPE das previsões dos modelos semanais e das previsões BMA para as seis semanas de antecedência.

A Figura 6.28 mostra as diferenças percentuais para o indicador NS, indicador de longo prazo, que avalia o poder de predição dos modelos hidrológicos. Para o índice NS, quanto maior o resultado, melhor, portanto os valores negativos (em azul) favorecem os resultados da aplicação da técnica BMA. Assim, o indicador NS também aponta os méritos da metodologia de mistura de modelos frente à modelagem semanal do ONS.

A Figura 6.29 traz as diferenças percentuais encontradas para o indicador DM, que é composto pelos indicadores NS e MAPE. Para o indicador DM, quanto mais próximos de zero os resultados, melhor. Assim, constata-se que em praticamente todos os casos houve melhora após as ponderações.

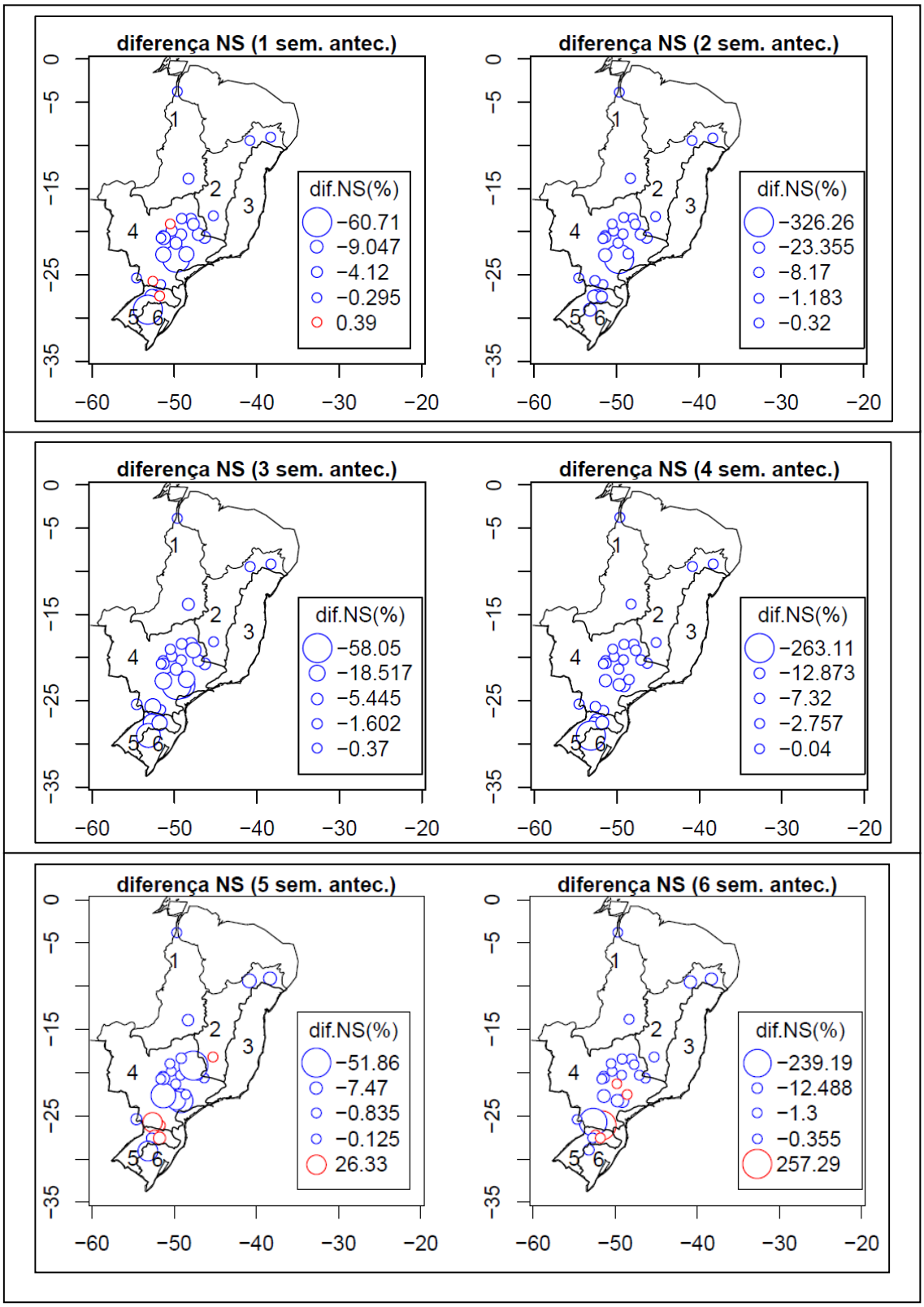

Figura 6.28 - Resultados das diferenças entre o indicador NS das previsões semanais e das previsões BMA para as seis semanas de antecedência.

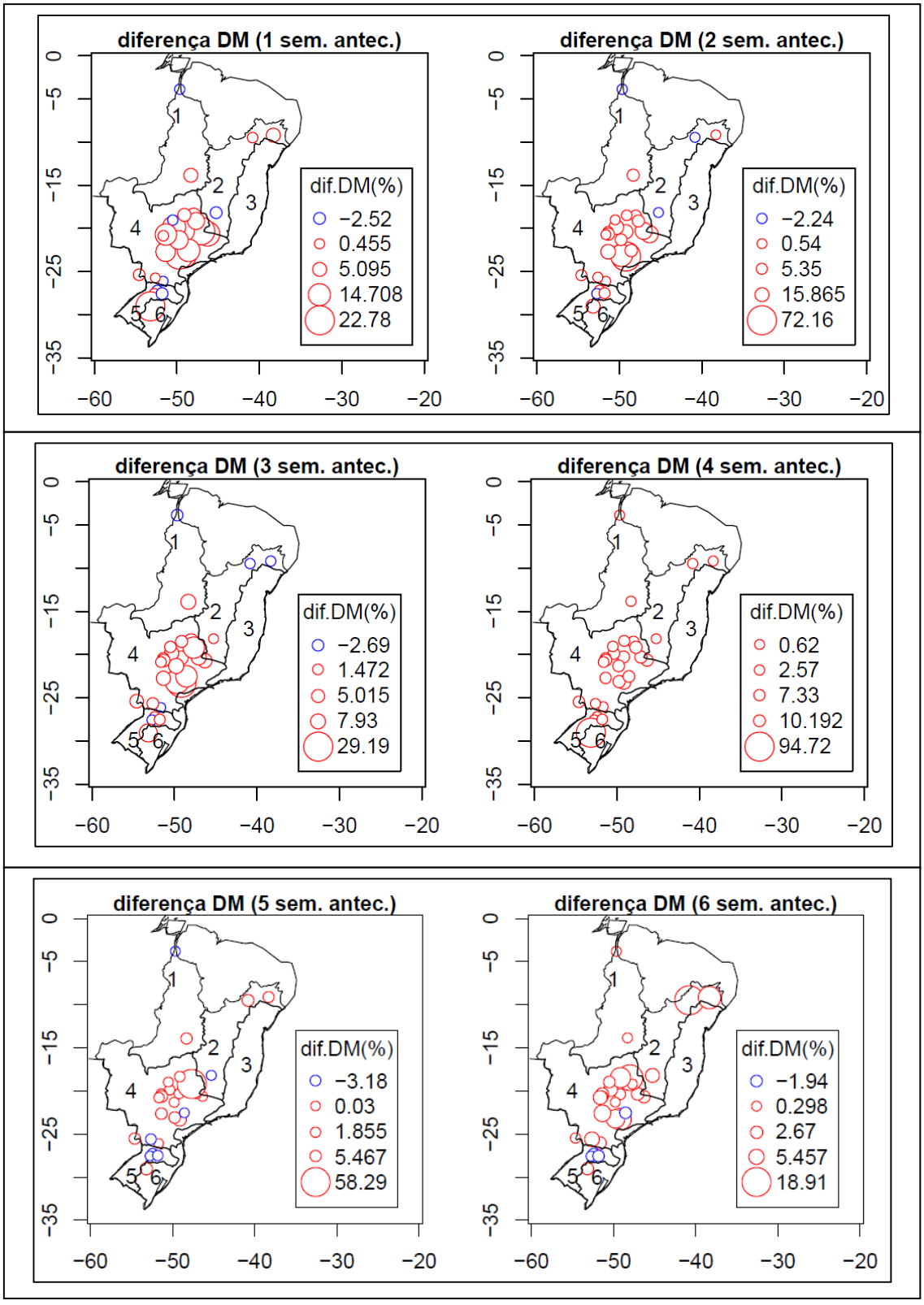

Figura 6.29 - Resultados das diferenças entre o indicador DM das previsões semanais realizadas pelo ONS e das previsões BMA de uma a seis semanas de antecedência.

6.3.3. Resultados localizados por bacias hidrográficas

O desempenho das previsões das usinas foi avaliado tanto individualmente como por bacia hidrográfica. Ao todo, dez bacias distintas participaram dos estudos, representadas pelos vinte e oito principais reservatórios de acumulação do país. Os resultados dos indicadores por usina encontram-se no Apêndice D desta dissertação.

A bacia do rio Paranaíba, por exemplo, contempla uma série de reservatórios em cascata, ilustrados na Figura 6.30. Nessa bacia, foram avaliadas as previsões de quatro de seus reservatórios: as UHEs Nova Ponte, Emborcação, Itumbiara e São Simão, com áreas de drenagem variando de 15.480 km^2, na primeira das usinas, a 171.474 km^2, na última delas.

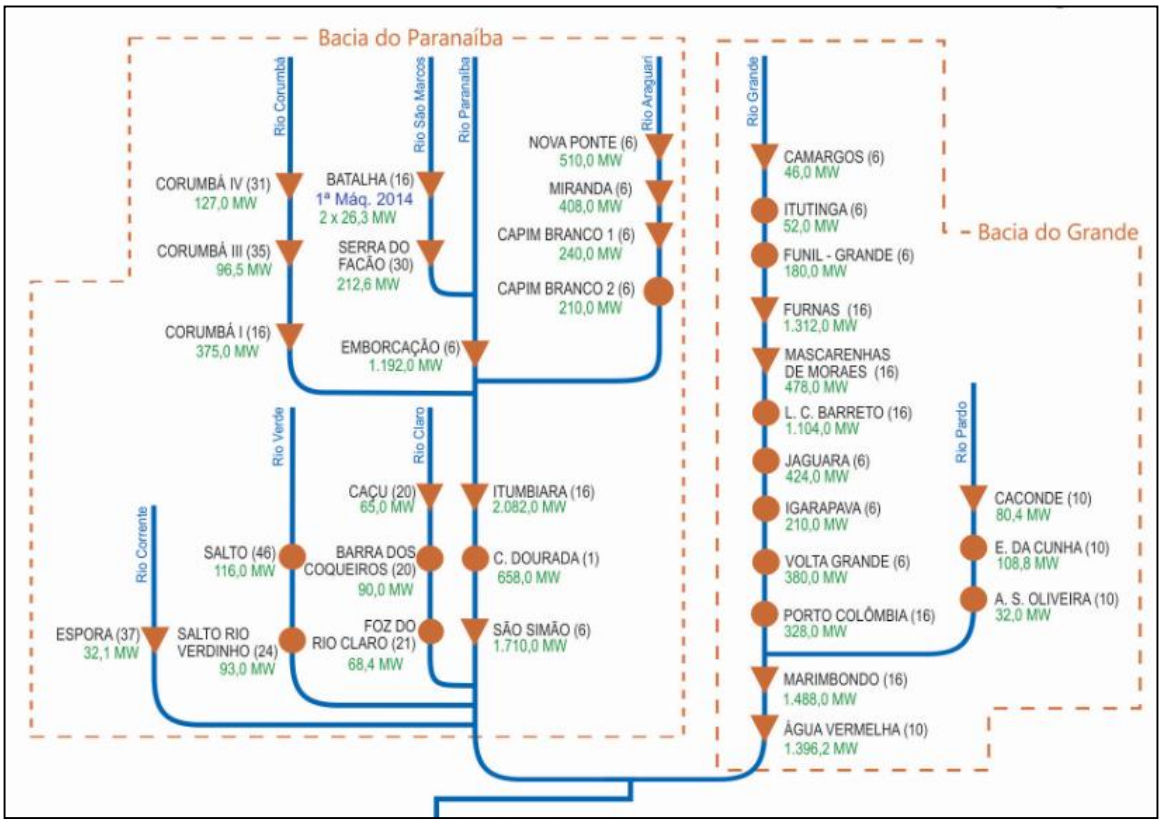

Figura 6.30 - Esquema topológico das usinas das bacias do Paranaíba e Grande.

Desempenho nos reservatórios da bacia do rio Paranaíba

Nas Figuras D.1 a D.4 do Apêndice D apresenta-se o comportamento dos indicadores das UHEs Nova Ponte, Emborcação, Itumbiara e São Simão. O painel superior da esquerda mostra a ponderação dos pesos semanais (curva na cor preta) e mensais (cor azul); o painel inferior da esquerda apresenta o indicador DM das previsões mensais (convertidos para a escala semanal); já o painel da direita, traz o indicador DM para as previsões semanais realizadas pelo ONS (curva na cor azul) e para as previsões BMA (na cor vermelha).

Na análise das quatro usinas, destaca-se que a maior diferença obtida entre o indicador DM das previsões semanais antes e após a aplicação da técnica BMA ocorreu nas UHEs Nova Ponte e Emborcação, situadas no trecho correspondente ao alto Paranaíba, e que utilizaram os modelos mensais com maior intensidade nas ponderações.

A Figura D.1 apresenta os indicadores para a UHE Nova Ponte. Verifica-se no painel da direita o grande descolamento entre as curvas azul e vermelha, o que demonstra o alto desempenho das ponderações BMA nessa usina, quando comparadas com as previsões originais. Nota-se, no painel superior da esquerda, que a utilização do peso mensal é crescente até a quinta semana, onde atinge seu pico (60% de utilização).

Quanto à UHE Emborcação, com área de drenagem de 29.050 km^2, os resultados de seus indicadores encontram-se na Figura D.2. Verifica-se que a superação do peso semanal pelo mensal ocorreu na transição da terceira para a quarta semana.

A Figura D.3 apresenta o resultado dos indicadores da UHE Itumbiara, que possui área de drenagem de 94.728 km^2, já no médio Paranaíba. Quanto à distribuição dos pesos, verifica-se que a superação do peso semanal pelo mensal ocorre na última semana, representando um deslocamento em relação ao ocorrido nas usinas avaliadas anteriormente. Em consequência, houve redução na diferença entre os indicadores BMA semanais antes e após as ponderações (menores diferenças entre a curva azul e vermelha no painel da direita).

Isso também pode ser verificado na Figura D.4, referente à UHE São Simão, situada no trecho do baixo Paranaíba, que apresentou a menor utilização do modelo mensal nas ponderações e o pior desempenho bayesiano em relação às demais usinas da cascata. Assim, quanto mais a montante, maior a utilização do modelo mensal e, portanto, melhores os resultados BMA. Quanto mais a jusante, menor o desempenho do modelo mensal nas

usinas da cascata do Paranaíba, reduzindo os pesos mensais e as diferenças entre os indicadores BMA semanais.

Pode-se dizer que as previsões dos modelos semanais adotados pelo ONS para o trecho alto da bacia do Paranaíba, com maior declividade, possuem maiores erros de previsão (maiores índices DM semanais – curvas na cor azul). Com isso, o modelo mensal é mais exigido nas ponderações, resultando em maiores reduções nos erros semanais (maior descolamento entre as curvas azul e vermelha no painel da direita das Figura D.1 a D.4). À medida que o rio alcança seus trechos médio e baixo, mais regularizados, de menor declividade e com maior área de drenagem (mais afluentes como contribuintes), melhoram-se as previsões semanais do ONS (menores DM semanais), reduzindo a utilização dos modelos mensais.

Desempenho nos reservatórios da bacia do rio Grande

Análise semelhante para as usinas avaliadas no rio Grande (UHEs Furnas, Mascarenhas, Maribondo e Água Vermelha, ilustradas no esquema topológico da Figura 6.30) aponta que, mesmo com ponderações mensais não superando os pesos semanais, os indicadores das usinas após a aplicação da técnica BMA igualaram ou superaram os indicadores dos modelos semanais do ONS. Exceção feita à UHE Furnas, na sexta semana, conforme Figura D.5.

Os gráficos das UHEs analisadas no rio Grande (Figuras D.5 a D.8) apresentam comportamento similar quanto à atribuição dos pesos aos modelos. Nota-se que o indicador DM mensal varia entre 0,45 e 0,55 em Furnas e Mascarenhas, apresentando os valores mais elevados das usinas. Esses valores reduzem à medida que o rio corre para jusante, como pode ser visto nas referidas figuras. As usinas de Marimbondo e de Água Vermelha, mais a jusante, apresentaram os melhores resultados da cascata, com os indicadores DM mensais entre 0,35 a 0,45. Além disso, os indicadores mensais apresentaram melhora no decorrer das semanas, estabilizando-se na quinta e sexta semana, resultados responsáveis pelos ganhos para as ponderações BMA semanais em todas antecedências para todos os reservatórios.

Desempenho nos reservatórios da bacia do rio Tietê

As usinas da bacia do rio Tietê que participaram dos estudos são: UHEs Barra Bonita, Promissão, Três Irmãos e Ilha Solteira, já na confluência com o rio Paraná. A UHE Barra Bonita apresentou comportamento interessante em relação aos pesos dos modelos, em que os modelos semanal e mensal tiveram usos bastante equilibrados até a quinta semana, com o modelo mensal sendo utilizado com maior intensidade entre a terceira e quarta semana, perdendo força nas duas últimas, graças ao baixo desempenho do modelo mensal neste período. Isto pode ser observado nos gráficos dos indicadores da Figura D.9, sendo que o indicador mensal obteve melhoras progressivas da primeira até a quarta semana, período em que o indicador DM_BMA obteve as melhores performances. A seguir, na Figura 6.31, o esquema topológico das bacias hidrográficas dos rios Tietê e Paranapanema.

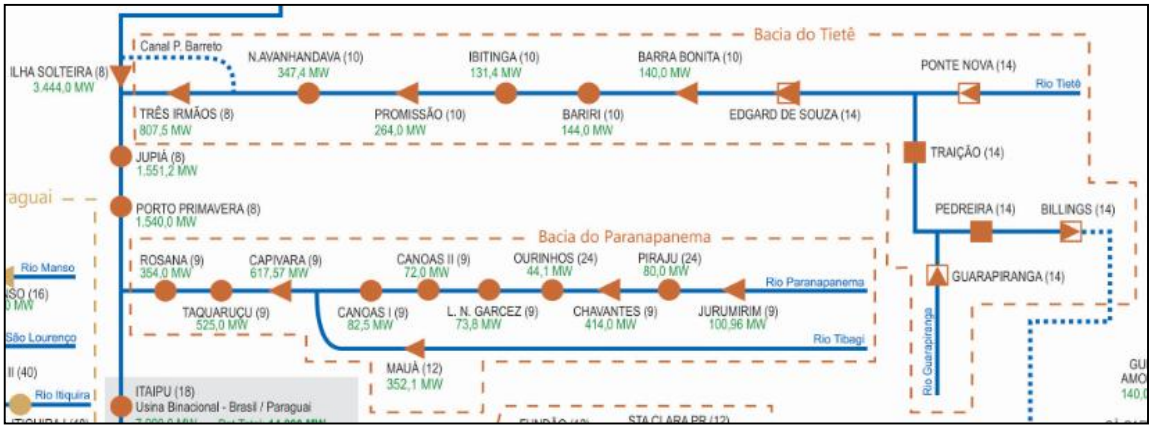

Figura 6.31 - Esquema topológico das usinas do Tietê-Paranapanema.

Conclusões similares às obtidas para a UHE Barra Bonita podem ser aplicadas às UHEs Promissão e Três Irmãos. Entretanto, algumas sutilezas marcam a composição dos pesos para estas usinas, como o fato de o modelo mensal ter maior peso que o semanal na quinta semana para a UHE Três Irmãos, o que propiciou um ganho de qualidade nas previsões BMA para a referida semana em relação às outras usinas.

Na Figura D.11 que contém os resultados da UHE Três Irmãos, nota-se uma tendência de recuperação na capacidade preditiva do modelo mensal da quinta para a sexta semana, o que refletiu num maior peso do modelo mensal para a UHE Três Irmãos nesta semana, quando comparada com as demais usinas do Tietê.

Quanto à UHE Ilha Solteira, por ter apresentado comportamento distinto das usinas do Tietê, tendo os aspectos da modelagem semelhantes aos obtidos na UHE Jupiá, sua análise ocorrerá em conjunto com a referida usina do rio Paraná.

Desempenho nos reservatórios da bacia do rio Parapanema

Em relação às usinas do rio Paranapanema, foram avaliadas as UHEs Jurumirim, Chavantes e Capivara. A UHE Jurumirim, também conhecida como Almirante A. Laydner apresentou resultados semanais positivos na aplicação da modelagem BMA quando comparados com as previsões semanais do ONS. Os pesos mensais mostraram-se relevantes desde a segunda semana, assumindo o protagonismo das ponderações a partir da quarta semana, o que pode ter contribuído para o desempenho BMA (Figura D.12).

Os resultados para as UHEs Chavantes e Capivara são semelhantes aos da UHE Jurumirim, com o desempenho dos modelos mensal e semanal um pouco inferiores. Como prováveis causas, o equilíbrio entre os pesos dos modelos na segunda e terceira semanas não ocorreu na mesma escala (Figura D.14). Além disso, na UHE Capivara não houve a preponderância do modelo mensal na quinta semana, como em Chavantes e em Jurumirim.

Desempenho nos reservatórios da bacia do rio Paraná

Na cascata do rio Paraná, constam das análises as UHE Ilha Solteira, Jupiá e Itaipu. As duas últimas não são reservatórios de acumulação, mas pela importância sistêmica, encontram-se entre os vinte e oito reservatórios com acompanhamento especial do ONS.

As UHEs Jupiá e Ilha Solteira alcançaram resultados semelhantes quando analisadas pelo quesito peso bayesiano, com baixos pesos advindos do modelo mensal. Mesmo recebendo como contribuintes as vazões do rio Tietê, que apresentou forte utilização dos pesos mensais, a UHE Jupiá obteve similaridade de resultados nas ponderações dos pesos com as usinas do rio Grande e do baixo Paranaíba, o mesmo ocorrendo com a UHE Ilha Solteira.

A Figura D.16 mostra os resultados da UHE Jupiá, que em nenhuma semana de antecedência teve os pesos semanais superados pelos mensais. Verifica-se que tanto o modelo mensal PARX quanto o modelo semanal do ONS apresentam desempenhos

satisfatórios em virtude dos baixos valores para o indicador DM, o que pode justificar a maior utilização do modelo semanal nas ponderações BMA. Mesmo assim, a técnica BMA apresentou melhorias nas previsões semanais a partir da segunda semana, graças aos bons resultados das previsões mensais.

Em relação à UHE Itaipu, usina de maior relevância para o Sistema Interligado Brasileiro, os resultados obtidos também apontam para uma melhoria nas previsões após a aplicação da técnica BMA a partir da segunda semana, com destaque entre a terceira e quinta semanas de antecedência. O modelo mensal de previsão mostrou-se estável nas seis semanas, contribuindo para os mencionados ganhos (Figura D.17).

Esses resultados são compatíveis com os obtidos por Lima (2014), que utilizou a abordagem BMA para acoplar as previsões de vazões mensais e diárias afluentes ao reservatório da UHE Itaipu. Em seu estudo, Lima concluiu que as previsões mensais de vazão começam a contribuir para as previsões na escala diária a partir de 7 dias de antecedência. Além disso, o modelo mensal assume o protagonismo nas previsões diárias do 16º dia em diante, sendo que desde o décimo dia, a ponderação BMA fornece melhorias significativas nas performances de previsão diária. Aqui, as previsões mensais melhoram as previsões na escala semanal a partir da segunda semana de antecedência.

Ainda no que se referem à bacia do Paraná, os rios que desembocam a jusante do reservatório de Itaipu são o Iguaçu, o Uruguai e o Paraguai. As usinas estudadas no rio Iguaçu foram as UHE Foz do Areia e Salto Santiago. A seguir, na Figura 6.32, ilustra-se o esquema topológico da bacia hidrográfica do rio Iguaçu.

Desempenho nos reservatórios da bacia do rio Iguaçu

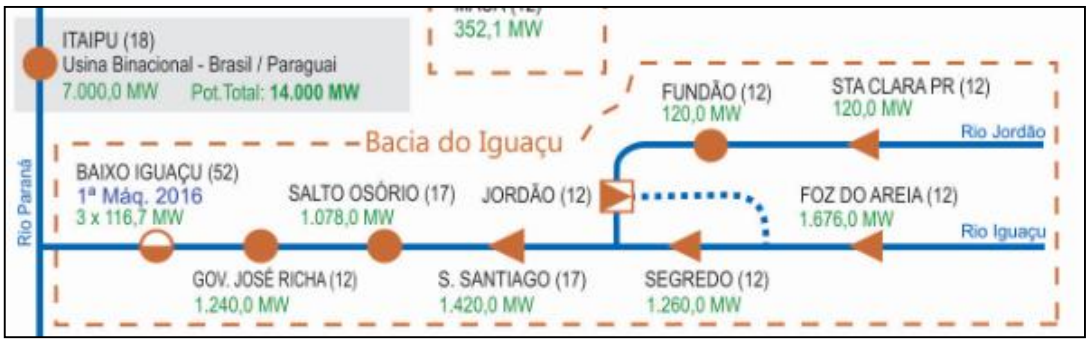

Figura 6.32 - Esquema topológico das usinas do rio Iguaçu.

Com exceção feita aos resultados da sexta semana, a UHE Foz do Areia (Governador Bento Munhoz), foi uma das poucas usinas que não tiveram suas previsões melhoradas com a aplicação da técnica BMA. O baixo desempenho dos modelos semanal e mensal (baixo indicador de qualidade NS e alto indicador MAPE) pode ter influenciado nos resultados. Da quinta para a sexta semana, registra-se que a melhoria verificada foi acompanhada de recuperação no desempenho no modelo mensal, apresentada no painel inferior à esquerda da Figura D.18 (que, diferentemente das figuras anteriores, ilustra o comportamento do indicador NS).

Quanto à UHE Salto Santiago, o comportamento dos resultados foi similar à da usina Foz do Areia até a terceira semana, quando o peso mensal passou a equilibrar-se com o semanal em termos de importância. Em consequência, melhores resultados foram obtidos nos modelos semanais com a técnica BMA nessa semanas (Figura D.19). Exceção obtida na quinta semana, coincidente com a repentina queda no indicador NS do modelo mensal.

Desempenho nos reservatórios da bacia dos rios Uruguai e Jacuí

O esquema topológico das bacias hidrográficas dos rios Uruguai e Jacuí encontram-se na Figura 6.33. Na bacia do rio Uruguai, fizeram parte dos estudos as UHEs Passo Fundo, Machadinho e Itá.

Os resultados na UHE Passo Fundo, situada em rio de mesmo nome, um dos afluentes do Uruguai, são apresentados na Figura D.20, onde aponta-se um equilíbrio entre as previsões semanais conduzidas pelo ONS antes e após a aplicação da ponderação bayesiana. Os baixos resultados dos indicadores semanais NS e MAPE indicam que o modelo semanal não é eficaz. Assim, mesmo com a indicação da boa utilização dos pesos mensais nas semanas intermediárias, a mistura de modelos não foi suficiente para propiciar melhorias para o período completo (2009 a 2014). Numa avaliação ano a ano, o modelo BMA obteve melhores resultados em 2009, 2010 e 2013, não sendo eficaz nos demais anos.

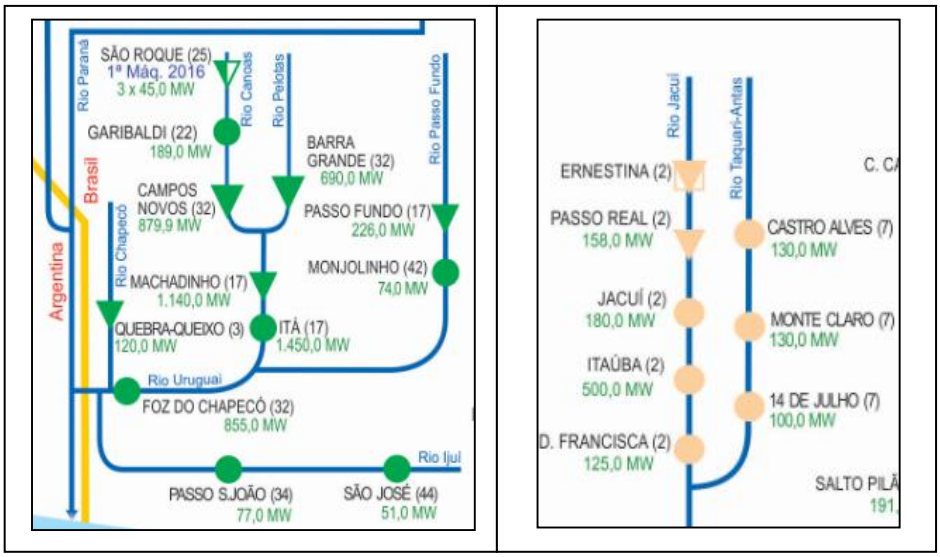

Figura 6.33 - Esquema topológico das usinas da bacia dos rios Uruguai e Jacuí.

As UHEs Itá e Machadinho, no rio Uruguai, obtiveram resultados similares entre si. Os destaques positivos do modelo BMA aconteceram entre a segunda e quarta semanas e os negativos, nas duas útimas semanas. Mesmo com a melhoria no indicador DM mensal no decorrer das semanas, os baixos resultados dos indicadores NS e MAPE semanais e mensais evitaram melhores desempenhos do modelo BMA, mesmo com os pesos mensais atingindo a soberania a partir da quarta semana. Ademais, registra-se que em avaliação complementar, realizada ano a ano, o modelo BMA obteve melhores resultados do que os aqui apresentados (resultados para todo o período). Na Figura D.22 resumem-se os resultados para a UHE Itá.

No rio Jacuí, fez parte das análises a UHE Passo Real, que obteve desempenho favorável após as ponderações dos modelos, quando comparado com o modelo semanal puro. Verifica-se que o baixo desempenho do modelo semanal entre a terceira e quinta semanas foi adequadamente sanado por meio da técnica BMA (Figura D.23).

Desempenho nos reservatórios da bacia dos rios Sobradinho e Tocantins

No rio Sobradinho, participaram dos estudos as UHEs Três Marias, Sobradinho e Itaparica; no Tocantins, as UHEs Serra da Mesa e Tucuruí. O esquema topológico das referidas bacias é apresentado na Figura 6.34.

A ponderação BMA resultou em melhorias para a UHE Três Marias apenas na sexta semana. Para a UHE Sobradinho, as melhorias ocorreram na quinta e sexta semanas. Destaque para os bons indicadores dos modelos mensais e semanais das usinas. Note que o indicador do modelo mensal da UHE Três Marias (Figura D.24) decresce com o decorrer das semanas, ocorrendo o inverso nas UHEs Sobradinho e Itaparica (Figuras D.25 e D.26). Conforme as análises anteriores nas demais usinas esperavam-se melhores resultados da aplicação da técnica BMA para as UHEs Sobradinho e Itaparica. Uma das prováveis justificativas reside na Tabela 6.4, que contem os resultados dos indicadores da UHE Sobradinho, retirados do Relatório Anual de Avaliação das Previsões de Vazões do ONS. Por utilizar o modelo de previsões de base física CPINS (propagação) nas primeiras duas semanas, as previsões nessa usina (e em Itaparica) possuem historicamente bons resultados. Assim, as melhorias, quando ocorrem, tendem a ser menores que em outras usinas.

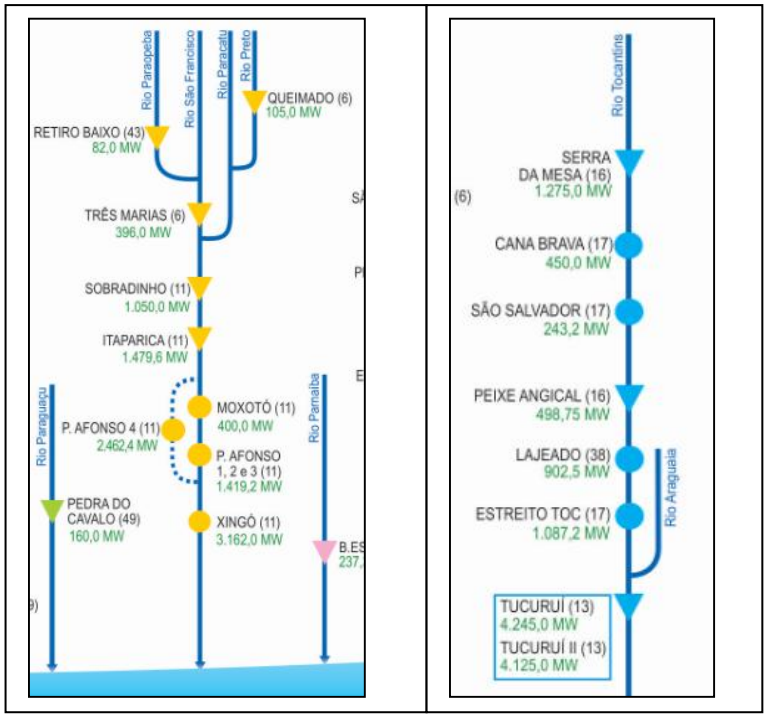

Figura 6.34 - Esquema topológico das usinas das bacias dos rios São Francisco e Tocantins.

Tabela 6.4 - Resultados para o indicador DM na UHE Sobradinho entre 2009 e 2013.

	Resultados para o indicador DM					
Antecedência	**2009**	**2010**	**2011**	**2012**	**2013**	**Valor médio**
1 semana	0,07	0,05	0,05	0,05	0,05	0,05
2 semanas	0,13	0,12	0,10	0,12	0,11	0,12
3 semanas	0,21	0,34	0,16	0,26	0,26	0,25

Na UHE Serra da Mesa, no rio Tocantins, os resultados BMA foram promissores em todas as semanas, com destaque para a segunda e terceira semana (Figura D.27).

Da mesma forma que nas usinas do rio São Francisco, os modelos mensal e semanal das UHEs do rio Tocantins também apresentaram bom desempenho, conforme Figura D.28. Diferente da UHE Serra da Mesa, em que o indicador DM mensal foi reduzindo com o decorrer das semanas, o indicador mensal da UHE Tucuruí apresentou crescimento a partir da segunda semana, o que pode ter contribuído para o desempenho semanal BMA um pouco abaixo do modelo semanal utilizado pelo ONS.

6.3.4. O teste estatístico de Diebold-Mariano

Com o intuito de atestar estatisticamente se as diferenças entre as competências dos modelos devem-se a erros de amostragem, foi testada a hipótese nula de que as saídas dos modelos semanais antes e após a ponderação BMA possuem iguais acurácias de precisão. A Figura 6.35 exemplifica os resultados dos testes unilaterais de Diebold-Mariano realizados para a UHE Tucuruí em função da antecedência semanal de previsões. Ao comparar os modelos, não se rejeita a hipótese alternativa de que o modelo BMA, ao nível de significância de 10%, possui maior acurácia de previsões que os modelos utilizados pelo ONS para antecedências de uma a cinco semanas (resultados dos p-valores abaixo da linha tracejada na cor verde, ou seja, inferiores a 0,10).

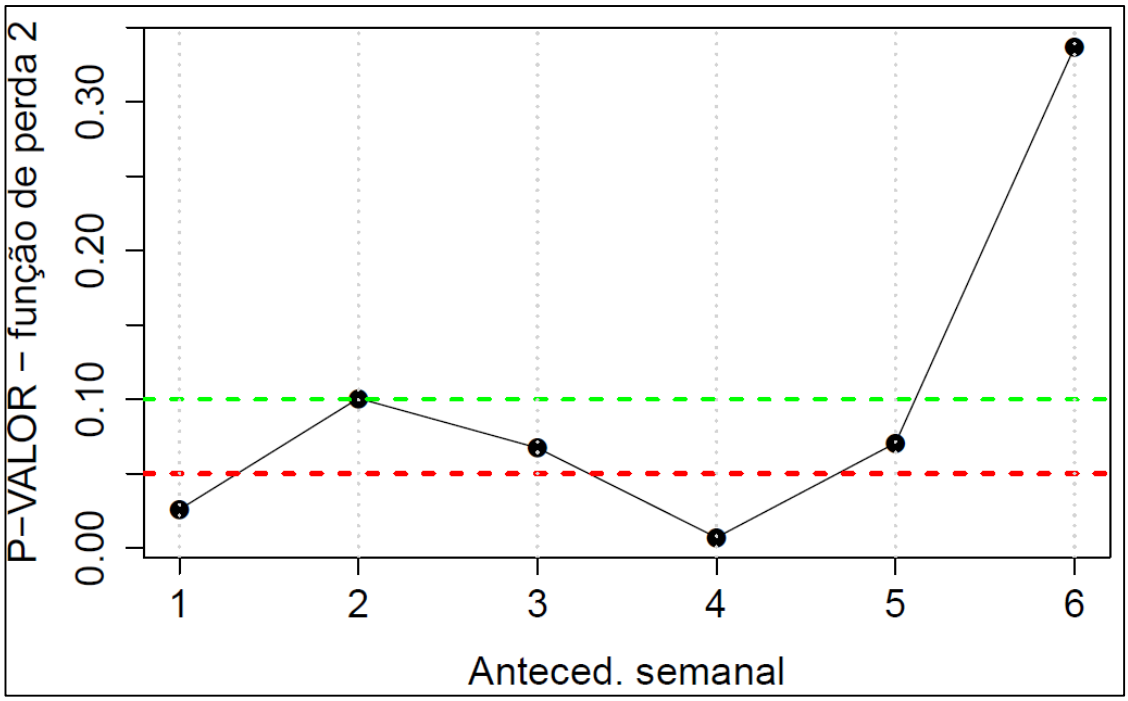

Figura 6.35 - Resultados do teste estatístico de Diebold-Mariano para a UHE Tucuruí, sob a hipótese nula de que os erros das previsões de vazões dos dois modelos são iguais. A linha horizontal em laranja é a referência para o p-valor = 0,05 e a linha verde, para o p-valor = 0,10.

O resultado dos testes de hipóteses para os vinte oito reservatórios encontra-se resumido na Tabela 6.5. Verifica-se que, em 102 das 168 previsões analisadas (seis antecedências semanais para cada um dos 28 reservatórios), rejeitou-se a hipótese nula, ao nível de significância de 10%. A discretização dos resultados de acordo com a antecedência semanal das previsões, efetuada na última linha da Tabela 6.5, mostra que os resultados são melhores nas quatro primeiras semanas, com índices de rejeição da hipótese nula variando entre 71,4% (primeira e quarta semana) e 82,1% (segunda semana de antecedência).

Tabela 6.5 - Resultados do teste de Diebold-Mariano para todos os vinte e oito reservatórios (casos de rejeição da hipótese nula, ao nível de significância de 10%).

Reservatório	Teste de Diebold-Mariano (P-VALOR 10%)						Bacia
ANTECEDÊNCIA SEMANAL	1	2	3	4	5	6	
FURNAS	0,00	0,01	0,01	0,02	0,47	0,16	Grande
MASCARENHAS DE MORAES	0,00	0,01	0,01	0,03	0,36	0,15	
MARIMBONDO	0,01	0,01	0,07	0,02	0,23	0,33	
ÁGUA VERMELHA	0,00	0,01	0,07	0,04	0,32	0,30	
EMBORCAÇÃO	0,01	0,05	0,03	0,07	0,14	0,09	Paranaíba
NOVA PONTE	0,03	0,05	0,02	0,10	0,05	0,42	
ITUMBIARA	0,07	0,07	0,04	0,04	0,06	0,11	
SÃO SIMÃO	0,97	0,23	0,09	0,09	0,26	0,24	
A.A. LAYDNER (JURUMIRIM)	0,01	0,02	0,02	0,01	0,01	0,07	Paranapanema
CHAVANTES	0,00	0,01	0,02	0,03	0,08	0,05	
CAPIVARA	0,00	0,01	0,06	0,01	0,05	0,05	
G.B. MUNHOZ (FOZ DO AREIA)	0,28	0,11	0,45	0,23	0,56	0,14	Iguaçu
SALTO SANTIAGO	0,53	0,04	0,13	0,31	0,77	0,09	
MACHADINHO	0,59	0,07	0,11	0,04	0,57	0,56	Uruguai
ITÁ	0,42	0,06	0,12	0,03	0,53	0,54	
PASSO FUNDO	0,01	0,05	0,05	0,23	0,42	0,33	
PASSO REAL	0,02	0,00	0,01	0,15	0,03	0,14	Jacuí
TRÊS MARIAS	0,21	0,26	0,15	0,48	0,82	0,15	São Francisco
SOBRADINHO	0,07	0,03	0,06	0,06	0,02	0,03	
ITAPARICA	0,02	0,01	0,05	0,08	0,06	0,08	
BARRA BONITA	0,00	0,01	0,02	0,02	0,49	0,82	Tietê
PROMISSÃO	0,00	0,04	0,02	0,01	0,42	0,55	
TRÊS IRMÃOS	0,00	0,03	0,06	0,02	0,14	0,42	
ILHA SOLTEIRA	0,33	0,28	0,20	0,15	0,28	0,17	Paraná
JUPIÁ	0,31	0,33	0,24	0,14	0,32	0,14	
ITAIPU	0,10	0,00	0,00	0,00	0,03	0,23	
SERRA DA MESA	0,02	0,01	0,03	0,12	0,10	0,21	Tocantins
TUCURUI	0,03	0,10	0,07	0,01	0,07	0,34	
rejeição da hipótese H0	71,4%	82,1%	75,0%	71,4%	39,3%	25,0%	

Resumo dos resultados

Por fim, a Tabela 6.6 compila os principais resultados obtidos em cada antecedência semanal de previsão para os vinte oito reservatórios, com destaque para a distribição dos pesos dos modelos e as tendências do indicador distância multicritério (DM) semanal e mensal. Considerando as 168 previsões analisadas (6 previsões para cada reservatório), a aplicação da técnica BMA superou as previsões semanais em 118 ocasiões,

ocorrendo 14 casos em que foi inferior e 36 casos de empate no desempenho das previsões. Se avaliadas apenas as previsões com três e quatro semanas de antecedência, a técnica BMA não superou ou foi similar às previsões semanais em apenas uma ocasião.

Em relação à distribuição dos pesos para os modelos, verifica-se forte relação entre a tendência do indicador DM mensal e a utilização dos pesos semanais nas ponderações, que foi ressaltada por meio da padronagem de cores adotada na Tabela 6.6. Quando o indicador DM possui tendência decrescente e/ou de estabilização na passagem de uma semana para outra (vide células nas cores azul clara e amarela, respectivamente), a importância do peso semanal acompanha tal tendência. Em outras palavras, a melhoria ou a manutenção no desempenho do modelo mensal no decorrer dos horizontes de previsão semanais das previsões, tende a resultar na redução do peso bayesiano semanal.

Da mesma forma, os modelos mensais que apresentaram queda de rendimento na transição dos resultados de lag mensal 1 para 2 foram responsáveis pela redução dos pesos bayesianos nessa escala de tempo, como pode ser visto na Tabela 6.6. Soma-se a isso, o fato de o número de amostras reduzir consideravelmente para os modelos com cinco e seis semanas de antecedência, o que pode ter contribuído para o menor desempenho das ponderações do que em outras antecedências. Essa condição foi objeto da análise de Gibbons et al. (2008), que concluíram que amostras de menor tamanho possuem menos informação nos dados, com isso os valores das probabilidades a *posteriori* dos modelos (PMPs) aproximam-se dos valores a priori, auferindo menores ganhos às ponderações que em modelos com maior número de amostras.

Quanto à preponderância dos pesos, doze reservatórios mantiveram o modelo semanal como o principal fornecedor de previsões. Nos demais, em três usinas o modelo mensal assumiu o protagonismo na segunda semana de antecedência; dois reservatórios, na terceira semana; seis, na quarta semana; um reservatório, na quinta semana e quatro, na sexta semana de antecedência.

Entretanto, verifica-se não ser necessário que o modelo mensal assuma o protagonismo das ponderações para que os resultados sejam alcançados. Haja vista a grande quantidade de previsões BMA que mostraram-se melhores ou similares aos resultados das previsões semanais realizadas pelo ONS, pode-se inferir que pequenas parcelas de contribuição dos modelos mensais nas ponderações, desde que atestado o bom desempenho dos modelos mensais, serão capazes de propiciar melhorias aos modelos semanais.

Tabela 6.6 - Resumo de alguns dos resultados por reservatório.

Bacia	Melhor desempenho: indicador DM semanal						Tendência do indicador DM Mensal (na escala semanal)						protagonismo do peso bayesiano						Tendência dos pesos bayesianos para o modelo semanal					Reservatório
	6	5	4	3	2	1	5 - 6	4 - 5	3- 4	2 - 3	1 - 2		6	5	4	3	2	1	5 - 6	4 - 5	3- 4	2 - 3	1 - 2	ANTECEDÊNCIA SEMANAL
Grande	BMA	empate	BMA	BMA	BMA	BMA	decresc.	crescente	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	decresc.	crescente	decresc.	crescente	decresc.	FURNAS
	BMA	empate	BMA	BMA	BMA	BMA	decresc.	crescente	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	estável	crescente	decresc.	crescente	decresc.	MASCARENHAS DE MORAES
	BMA	BMA	BMA	BMA	BMA	BMA	estável	crescente	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	estável	crescente	decresc.	estável	decresc.	MARIMBONDO
	BMA	BMA	BMA	BMA	BMA	BMA	crescente	crescente	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	decresc.	crescente	decresc.	estável	decresc.	ÁGUA VERMELHA
Paranaíba	BMA	BMA	BMA	BMA	BMA	BMA	decresc.	decresc.	decresc.	decresc.	decresc.		MÊS	MÊS	MÊS	SEM	SEM	SEM	decresc.	crescente	decresc.	decresc.	decresc.	EMBORCAÇÃO
	empate	BMA	BMA	BMA	BMA	BMA	estável	decresc.	decresc.	decresc.	decresc.		SEM	MÊS	SEM	SEM	SEM	SEM	crescente	decresc.	decresc.	decresc.	decresc.	NOVA PONTE
	BMA	BMA	BMA	BMA	BMA	BMA	decresc.	decresc.	decresc.	decresc.	decresc.		MÊS	SEM	SEM	SEM	SEM	SEM	decresc.	crescente	decresc.	decresc.	decresc.	ITUMBIARA
	BMA	empate	BMA	BMA	empate	empate	decresc.	estável	estável	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	decresc.	crescente	decresc.	decresc.	decresc.	SÃO SIMÃO
Paranapanema	BMA	BMA	BMA	BMA	BMA	BMA	crescente	crescente	estável	estável	estável		empate	MÊS	MÊS	SEM	SEM	SEM	crescente	decresc.	decresc.	estável	decresc.	A.A. LAYDNER (JURUMIRIM)
	BMA	BMA	BMA	BMA	BMA	BMA	estável	crescente	estável	estável	estável		SEM	MÊS	MÊS	SEM	SEM	SEM	crescente	decresc.	decresc.	estável	decresc.	CHAVANTES
	BMA	BMA	BMA	BMA	BMA	BMA	estável	crescente	estável	estável	estável		SEM	SEM	MÊS	SEM	SEM	SEM	estável	crescente	decresc.	decresc.	decresc.	CAPIVARA
Iguaçu	BMA	empate	BMA	empate	empate	empate	decresc.	crescente	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	decresc.	crescente	decresc.	estável	estável	G.B. MUNHOZ (FOZ DO AREIA)
	BMA	ONS	BMA	BMA	BMA	empate	decresc.	crescente	decresc.	decresc.	estável		MÊS	SEM	empate	SEM	SEM	SEM	estável	estável	decresc.	decresc.	estável	SALTO SANTIAGO
Uruguai	ONS	ONS	BMA	BMA	BMA	ONS	decresc.	estável	decresc.	decresc.	decresc.		MÊS	MÊS	MÊS	SEM	SEM	SEM	crescente	crescente	decresc.	decresc.	decresc.	MACHADINHO
	ONS	ONS	BMA	BMA	BMA	ONS	decresc.	estável	decresc.	decresc.	decresc.		MÊS	MÊS	MÊS	SEM	SEM	SEM	crescente	crescente	decresc.	decresc.	decresc.	ITÁ
	empate	empate	BMA	empate	ONS	empate	decresc.	estável	decresc.	decresc.	decresc.		SEM	SEM	SEM	MÊS	MÊS	SEM	crescente	crescente	crescente	decresc.	decresc.	PASSO FUNDO
Jacuí	BMA	BMA	BMA	BMA	BMA	BMA	decresc.	estável	estável	decresc.	decresc.		SEM	MÊS	MÊS	MÊS	MÊS	SEM	estável	crescente	decresc.	estável	decresc.	PASSO REAL
São Francisco	BMA	ONS	empate	empate	empate	ONS	decresc.	estável	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	decresc.	estável	estável	estável	estável	TRÊS MARIAS
	BMA	BMA	empate	empate	empate	empate	estável	crescente	estável	estável	estável		MÊS	SEM	SEM	SEM	SEM	SEM	decresc.	decresc.	decresc.	decresc.	decresc.	SOBRADINHO
	BMA	BMA	empate	empate	empate	empate	estável	crescente	estável	estável	estável		MÊS	SEM	SEM	SEM	SEM	SEM	decresc.	decresc.	decresc.	decresc.	decresc.	ITAPARICA
Tietê	ONS	ONS	BMA	BMA	BMA	BMA	crescente	crescente	decresc.	decresc.	decresc.		SEM	SEM	MÊS	MÊS	SEM	SEM	crescente	crescente	estável	decresc.	decresc.	BARRA BONITA
	empate	BMA	BMA	BMA	BMA	BMA	crescente	crescente	decresc.	decresc.	decresc.		SEM	SEM	MÊS	MÊS	SEM	SEM	crescente	crescente	decresc.	decresc.	decresc.	PROMISSÃO
	BMA	BMA	BMA	BMA	BMA	BMA	decresc.	crescente	decresc.	decresc.	decresc.		SEM	MÊS	MÊS	SEM	SEM	SEM	crescente	crescente	decresc.	estável	decresc.	TRÊS IRMÃOS
Paraná	BMA	BMA	BMA	BMA	BMA	empate	estável	estável	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	crescente	crescente	estável	decresc.	decresc.	ILHA SOLTEIRA
	BMA	BMA	BMA	BMA	empate	empate	estável	estável	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	decresc.	crescente	decresc.	decresc.	decresc.	JUPIÁ
	empate	BMA	BMA	BMA	BMA	empate	decresc.	crescente	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	crescente	crescente	decresc.	decresc.	decresc.	ITAIPU
Tocantins	empate	BMA	BMA	BMA	BMA	BMA	decresc.	estável	decresc.	decresc.	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	estável	estável	crescente	decresc.	decresc.	SERRA DA MESA
	empate	empate	BMA	ONS	ONS	empate	estável	crescente	estável	estável	decresc.		SEM	SEM	SEM	SEM	SEM	SEM	crescente	crescente	estável	estável	decresc.	TUCURUI

7. CONCLUSÕES E RECOMENDAÇÕES

Neste trabalho formulou-se uma proposta para o aperfeiçoamento das previsões de afluências semanais aos reservatórios do sistema hidroelétrico brasileiro. Os modelos atuais de previsão e simulação de afluências para as diversas escalas temporais (diária, semanal, mensal) utilizados pelo Operador Nacional do Sistema Elétrico (ONS) tendem a ser limitados no horizonte de previsão e na modelagem da estrutura de dependência existente entre as diversas escalas, reduzindo a qualidade das previsões e simulações de afluências.

As previsões mensais obtidas do modelo periódico auto-regressivo exógeno (PARX), que considera a informação climática e implicitamente a variabilidade espaço-temporal das séries históricas de afluências, foram acopladas às previsões semanais dos modelos utilizados pelo ONS utilizando-se a técnica de ponderação bayesiana de modelos (BMA) com o intuito de melhorar as previsões de vazões de curto prazo, particularmente as semanais e em horizontes de previsão que vão além da persistência da vazão (da segunda semana em diante).

Os resultados apontam para o êxito da aplicação da metodologia BMA em praticamente todos os vinte e oito principais reservatórios de acumulação, comprovados pela melhoria dos indicadores de qualidade adotados. O acoplamento bayesiano das previsões mensais de vazão começa a fornecer informação útil para as previsões semanais mesmo a partir das previsões com uma semana de antecedência, tornando-se inclusive o modelo preferencial para previsões de alguns dos reservatórios a partir da segunda semana, como foi o caso das UHEs Jurumirim (A.A. Laydner), Passo Fundo e Passo Real.

A avaliação das distribuições dos pesos dos modelos apresentou o gradativo crescimento da importância do modelo mensal, que atingiu o protagonismo dos modelos em horizontes de previsão semanais distintos para dezesseis reservatórios: em três reservatórios, o modelo assumiu o protagonismo na segunda semana; na terceira semana, duas usinas; cinco, na quarta semana; duas usinas, na quinta e quatro, na sexta semana de antecedência.

Dos vinte e oito reservatórios estudados, doze elegeram a atualização das previsões semanais a partir da modelagem proposta como o principal fornecedor de previsões em todas as semanas: as quatro usinas do rio Grande, além da UHE São Simão, ainda no rio Paranaíba, mas próximo da confluência com o Grande, bem como para a

sequência de usinas da cascata do Paraná (Ilha Solteira, Jupiá e Itaipu). Além das oito usinas situadas nos rios Grande e Paraná, o mesmo foi constatado para as UHEs Serra da Mesa e Tucuruí (rio Tocantins), a UHE Três Marias (rio São Francisco) e a UHE Foz do Areia (rio Iguaçu).

Além disso, verificou-se uma tendência de redução no desempenho do modelo PARX mensal na transição do primeiro para o segundo mês para os reservatórios do rio Grande, Paranapanema, Tietê e Iguaçu, sendo que na usinas do rio Iguaçu ocorreu uma recuperação no desempenho de previsão da quinta para a sexta semana. A referida recuperação repetiu-se na UHE Três Irmãos, a última usina da cascata do rio Tietê, próxima à confluência com o Paraná, e na UHE Itaipu. Essa capacidade de recuperação foi observada tanto pelos indicadores de qualidade mensais quanto pelos indicadores da aplicação do BMA.

Como reflexo direto da redução de desempenho do modelo PARX da quarta para a quinta semana, que em alguns casos foi mantida na sexta semana, constatou-se a redução dos pesos do modelo mensal nos reservatórios acima mencionados, principalmente nos rios Grande e Paranapanema, e novamente nas UHE Itaipu, Três Irmãos e São Simão.

Assim, verifica-se haver forte relação entre a tendência do indicador DM mensal e a utilização dos pesos semanais nas ponderações, quanto ao critério de distribuição dos pesos dos modelos. Quando o indicador DM possui tendência decrescente e/ou de estabilização na passagem de uma antecedência semanal para outra, a importância do peso semanal acompanha tal tendência.

Ainda em relação aos pesos, constatou-se em alguns casos a variação dos pesos dentro de uma mesma bacia, como o exemplo do rio Paranaíba, em que as previsões para o trecho alto do rio, com maior declividade, utilizaram-se mais do modelo mensal nas ponderações. À medida que o rio flui para seus trechos médio e baixo, reduz-se a utilização dos modelos mensais. Quanto maior a utilização dos pesos mensais, maiores as melhorias nas previsões semanais.

Dentro da lógica mencionada acima, os resultados para as UHEs Barra Bonita, Promissão e Três Irmãos, situadas na cascata do rio Tietê, mostram que o peso do modelo mensal supera o peso semanal na terceira e quarta semana, para as duas primeiras usinas, e na quarta e quinta semana, para a terceira das UHEs. A partir dos gráficos dos indicadores de qualidade, o desempenho das previsões semanais do ONS decresce de montante para jusante. Após a ponderação BMA, todas as três usinas obtiveram melhores previsões, para

todas as antecedências semanais, sendo que a primeira das três usinas teve maiores melhorias nas três primeiras semanas; a segunda usina, nas quatro primeiras e a última, na primeira, segunda, quarta e quinta semanas. No caso da cascata do rio Tietê, conclui-se que o deslocamento nas distribuições dos pesos dos modelos no sentido montante-jusante, no mesmo sentido do fluxo da bacia, pode ter influenciado no deslocamento dos resultados das ponderações BMA. Os melhores resultados BMA também decrescem no sentido montante-jusante.

Quanto às usinas cujo modelo mensal não se impôs ao semanal, verifica-se que há casos, como os das UHEs Ilha Solteira, Jupiá e Itaipu, em que tanto o modelo mensal quanto o modelo semanal apresentam bom desempenho, justificando a não preponderância do modelo mensal nas ponderações BMA, ou seja, maior utilização do modelo semanal. Mesmo nesses casos, a ferramenta BMA apresentou-se capaz de melhorar as previsões semanais. As melhorias ocorreram a partir da segunda semana, para as UHEs Jupiá e Itaipu, e a partir da terceira semana, para a UHE Ilha Solteira.

Haja vista a grande quantidade de previsões BMA que mostraram-se melhores ou similares aos resultados das previsões semanais realizadas pelo ONS, verifica-se não ser necessário que o modelo mensal assuma o protagonismo das ponderações para que os resultados sejam alcançados. Pequenas parcelas de contribuição dos modelos mensais nas ponderações, desde que atestado o bom desempenho dos modelos mensais, serão capazes de propiciar melhorias aos modelos semanais.

A maior competência do modelo BMA foi avaliada estatisticamente por meio do teste de Diebold-Mariano, onde conclui-se que 60,7% das 168 previsões são estatisticamente significantes, ao nível de 10%. Se consideradas apenas as quatro primeiras semanas de antecedência nas previsões, a taxa de aceitação a favor do modelo BMA em detrimento às previsões do ONS eleva-se, variando entre 71,4% e 82,1%.

Assim sendo, resta-se comprovada que a metodologia BMA, ao acoplar previsões mensais aos resultados das previsões semanais, é uma ferramenta eficaz ao aperfeiçoamento das previsões semanais de afluências aos reservatórios do SIN.

Por fim, apresentam-se a seguir algumas recomendações para futuros trabalhos:

- Utilizar os resultados obtidos nas previsões em modelos energéticos para quantificar os ganhos em termos de energia (ou em termos de redução de custos, emissão de gases de efeito estufa, etc.);

- Acoplar os resultados obtidos das simulações nas usinas de montante para as ponderações bayesianas da usina de jusante, buscando ampliar a correlação espacial entre elas;
- Avaliar melhorias no modelo mensal de previsão de afluências a partir do uso de outros indicadores climáticos, por exemplo, os relacionados aos eventos El Niño Modoki e indicadores de base trimestral;
- Avaliar o uso de indicadores climáticos previstos por modelos numéricos de tempo.

8. REFERÊNCIAS BIBLIOGRÁFICAS

Adamowski, J. F.; Khosa, R.; Rathinasamy, M. (2013) Multiscale Stream flow Forecasting Using a New Bayesian Model Average Based *Ensemble* Multi-Wavelet Volterra Nonlinear Method, American Geophysical Union, Fall Meeting 2013, Dezembro 2013.

Anderson, M. L.; Kavvas, M. L.; Mierzwa M. D. (2001) Probabilistic/*ensemble* forecasting: A case study using hydrologic response distributions associated with El Niño/Southern Oscillation (ENSO), Journal of Hydrology, 249, páginas 134-147, 2001.

Araghinejad, S.; Bur, D. H.; Karamouz, M. (2006) Long-*lead* probabilistic forecasting of streamflow using oceanatmospheric and hydrological predictors. Water Resources Research, vol. 42, 2006.

Box, G.E.P.; Cox, D.R. (1964) An analysis of transformations. Journal of the Royal Statistical Society, B.26, p.211-243, discussion 244-252.

Box, G.E.P.; Jenkins, G.M. (1970) Time series analysis-forecasting and control, Holden-Day.

Brockwell, P. J.; Davis, R. A. (2002) Introduction to Time Series and Forecasting, Springer, Segunda edição, 2002.

Castanharo, G.; Gibertoni, R.F.C; Müller, I.I.; Andriolo, M.V.; Kaviski, E.; Guilhon, L.G.F.; Rocha, V.F. (2007) Previsão de vazões na bacia do rio Iguaçu baseada no modelo SMAP e com incorporação de informações de precipitação. RBRH, vol.12, nº3, JUL/SET 2007, p. 57-68.

Cataldi, M.; Machado, C.O.; Guilhon, L.G.F.; Chou, S.C.; Gomes, J.L.; Bustamante, J.F. (2007) Análise das previsões de precipitação obtidas com a utilização do modelo Eta como insumo para modelos de previsão semanal de vazão natural. RBRH, vol.12, nº3, JUL/SET 2007, p. 5-12.

Cavalcanti, I. F. A.; Ferreira, N. J.; Silva, M. A. A. J.; Dias, M. A. F. S. [organizadores] (2009), Tempo e Clima no Brasil. Oficina de Textos, são Paulo, 2009.

Citeau, J.; Berges, J.; Demarcq, H.; Mahe, G. (1988) The watch of ITCZ migrations over tropical Atlantic as an indicator in drought forecast over Sahelian area. Ocean-Atmosphere Newsletter, v. 45 p. 1-3, 1988.

Climate Prediction Center. Frequently asked questions about El Niño and La Niña. http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensofaq.shtml. Acesso em 05/12/2014.

Collischonn, W.; Tucci, C.E.M.; Clarke, R.T.; Corbo, M.D.; Silva, B.C.; Collischonn, B.; Allasia, D.; Paz, A.R. (2007) Modelo hidrológico distribuído para previsão de vazões incrementais na bacia do Rio Paranaíba entre Itumbiara e São Simão. RBRH, vol.12, nº3, JUL/SET 2007, p. 43-55.

Collischonn, W.; Tucci, C.E.M. (2001) Simulação hidrológica de grandes bacias. Revista Brasileira de Recursos Hídricos, v. 6, n. 1, 2001.

Corps of Engineers, North Pacific Division (1967) Draft of Operating Instructions, Computer Program for Streamflow Synthesis and Reservoir Regulation, Portland, Oregon – USA.

Costa, F. S.; Damázio, J.M.; Maceira, M. E. P.; Souza, M.D.S.; Guilhon, L. G. E; Silva, S. B. (2003) Modelagem estocástica de previsão de vazões mensais: PREVIVAZM. XV Simpósio Brasileiro de Recursos Hídricos, Curitiba, PR, Brasil.

Costa, F.S.C.; Maceira, M.E.P.; Damázio, J.M. (2007) Modelos de previsão hidrológica aplicados ao planejamento da Operação do Sistema Elétrico Brasileiro. RBRH, vol.12, nº3, JUL/SET 2007, p. 21-30.

CPTEC – Centro de Previsão de Tempo e Estudos Climáticos. El Niño e La Niña. Disponível em: <http://www.cptec.inpe.br/enos>. Acesso em 05/12/2014.

Dempster, A.; Laird, N.; Rubin, D. (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion), Journal of the Royal Statistical Society Series B 39: 1–38. 1977.

Diebold, F.X.; Mariano, R. (1995) Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–265, 1995.

Do, C. B.; Batzoglou, S. (2008) What is the expectation maximization algorithm? Nature Biotechnology, Volume 26, número 8, p. 897-899, Agosto 2008.

Dong, L; Xiong, L.; Yu, K. (2013) Uncertainty Analysis of Multiple Hydrologic Models Using the Bayesian Model Averaging Method, Journal of Applied Mathematics,Volume 2013, Article ID 346045, 2013.

Duan, Q.; Ajami, N. K.; Gao, X.; Sorooshian, S. (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging, Advances in Water Resources, 30 (2007), 1371–1386, 2007.

Ehlers, R. (2014) Notas de Aula Análise de Inferência Bayesiana. http://www.icmc.usp.br/pessoas/ehlers/bayes/. Acesso em 05/12/2014.

Evsukoff, A.G; Ebecken, N.F.F. (2004) Identification of recurrent fuzzy systems with genetic algorithms. Proceedings of IEEE International Conference on Fuzzy Systems, pp. 1703-1708, vol.3, 2004.

Evsukoff, A.G; Ebecken, N.F.F.; Souza, F.T.; Alegre, M.P.; Terra, G.S.; Tavares, G.M.; Hora, A.F. (2007) Uma abordagem de mineração de dados para a previsão de vazões com incorporação de previsão de precipitação da bacia do Rio Iguaçu. Seminário de Previsão de Vazões do ONS, 2007.

Fernandez, C.; Ley, E.; Steel, M. F. J. (2001) Model uncertainty in cross-country growth regressions. Journal of Applied Econometrics 16:563–576, 2001.

Figueiredo, K.; Barbosa, C.R.H; Cruz, A.V.A.; Vellasco, M. M. B. R.; Pacheco, M.A.C.; Contreras, R.J; Barros, M.; Souza, R.C.; Marques, V.S.; Duarte, U.M.; Mendes, M.H. (2007) Modelo de previsão de vazão com informação de precipitação utilizando redes neurais. RBRH, vol.12, nº3, JUL/SET 2007, p. 69-82.

Fraley, C.; Raftery, A. E.; Gneiting, T.; Sloughter, J.M. (2010) *ensembleBMA*: Probabilistic forecasting using *ensembles* and Bayesian Model Averaging. URL http://CRAN.R-project.org/package=*ensembleBMA*. R package version 4.5., 2010.

Garcia, A. G. N. (2005) Representação do processo estocástico de energias afluentes por modelos auto-regressivos periódicos no planejamento de sistemas hidrotérmicos. Dissertação de Mestrado, Universidade Federal de Santa Catarina, Departamento de Engenharia Elétrica, Florianópolis-SC, 138p, outubro de 2005.

Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. (2003) Bayesian Data Analysis. Chapman & Hall/CRC Texts in Statistical Science. 2003.

Gelman, A.; Hill, J. (2007) Data Analysis Using Regression and Multilevel/Hierarchical Models Cambridge University Press, 2007.

Gibbons, J.M.; Cox, G.M.; Wood, A.T.A.; Craigon, J.; Ramsden, S.J.; Tarsitano, D.; Crout, N.M.J. (2008) Applying Bayesian Model Averaging to mechanistic models: An example and comparison of methods. Environmental Modelling & Software 23 (2008) 973-985, 2008.

Gneiting, T.; Raftery, A. E. (2005) Weather forecasting with *ensemble* methods. Science, 310:248–249, 2005.

Gneiting, T.; Balabdaoui, F.; Raftery, A. E. (2007) Probabilistic forecasts, calibration, and sharpness. Journal of the Royal Statistical Society, Series B, 69:243–268, 2007.

Grimm, A.M. (2004) How do La Niña events disturb the summer monsoon system in Brazil? Climate Dynamics 22, 123–138, 2004.

Guilhon, L.G.F. (2003) Modelo heurístico de previsão de vazões naturais médias semanais aplicado à Usina de Foz do Areia. Tese de Mestrado, UFRJ, 2003.

Guilhon, L.G.F.; Rocha, V.F.; Moreira, J.C. (2007) Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. RBRH, vol.12, nº3, JUL/SET 2007, p. 13-20.

Hamlet, A. F.; D. P. Lettenmaier (1999) Columbia River streamflow forecasting based on ENSO and PDO climate signals. Journal of Water Resources Planning Management, 125(6), p. 333– 341, 1999.

Hastenrath, S.; L. Heller (1977) Dynamics of climate hazards in northeast Brazil. Quarterly Journal of the Royal Meteorological Society v. 103, p. 77-92, 1977.

Hastie, T. J.; Tibshirani, R. J.; Friedman, J. (2008) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer-Verlag. Segunda Edição, 2008.

Hemri, S.; Fundel, F.; Zappa, M. (2013) Simultaneous calibration of ensemble river flow predictions over an entire range of lead times, Water Resources Research, Volume 49, p. 6744 - 6755, Outubro 2013.

Hoeting, J. A.; Madigan, D.; Raftery, A. E.; Volinsky, C. T. (1999) Bayesian Model Averaging: A Tutorial. Statistical Science 14:382–401, 1999.

Hurst, H. (1957) A suggested statistical model for some time series that occur in nature. Nature 180:494

Ilich, N.; Despotovic, J. (2008) A simple method for effective multi-site generation of stochastic hydrologic time series. Stoch Environ Res Risk Assess 22, páginas 265–279, 2008.

Jiménez, K.Q.; Collischonn, W. (2015) Método de combinação de dados de precipitação estimados por satélite e medidos em pluviômetros para a modelagem hidrológica. Revista Brasileira de Recursos Hídricos, v. 20, n. 1, 2015.

Kaplan, A.; Cane, M.; Kushnir, Y.; Clement, A.; Blumenthal, M.; Rajagopalan, B. (1998) Analyses of global sea surface temperature 1856–1991. Journal of Geophysical Research 103 (18), 518–567. 589, 1998.

Kelman J.; Damázio J.M.; Costa J.P. (1983) Geração de séries sintéticas de vazões diárias – Modelo Diana, Revista Brasileira de Engenharia – Caderno Recursos Hídricos, vol. 1, no 2, Brasil.

Kelman, J.; Vieira, A. M.; Rodriguez-Amaya, J. E. (2000) El Niño influence on streamflow forecasting. Stochastic Environmental Research and Risk Assessment 14, páginas 123-138, 2000.

Knashawn, H. M.; Ibrahim, J. G.; Chen, C.; Ryan, L. M. (2006) Bayesian Model Averaging With Applications to Benchmark Dose Estimation for Arsenic in Drinking Water. Journal of the American Statistical Association 101:9–17, 2006.

Leamer, E. (1978) Specification Searches: Ad Hoc Inference with Nonexperimental Data. New York: Wiley, 1978.

Liang, Z.; Wang, D.; Guo, Y.; Zhang, Y.; Dai, R. (2013) Application of Bayesian Model Averaging Approach to Multimodel *Ensemble* Hydrologic Forecasting, Journal of Hydrologic Engineering, 18(11), p. 1426–1436, 2013.

Lima, C. H. R.; Lall, U. (2008a) Inclusão de informação climática e variabilidade espacial em modelos estatíscos de previsão de afluências. II Simpósio de Recursos Hídricos do Sul-Sudeste, Rio de Janeiro, 2008.

Lima, C. H. R.; Lall, U. (2008b) Hidroclimatologia de cheias: eventos extremos, não-estacionariedade, processos climáticos de larga escala e uma perspectiva para simulações

num contexto de mudanças climáticas globais. II Simpósio de Recursos Hídricos do Sul Sudeste, Rio de Janeiro, 2008.

Lima, C. H. R.; Lall, U. (2010) Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model. Journal of Hydrology (Amsterdam) v. 380, p. 438-449, 2010.

Lima, C. H. R. (2014) Connecting Multiscale Streamflow Forecasts for the Hydropower System in Brazil Using Bayesian Model Averaging and Markov Chain Monte Carlo Sampling, ITISE 2014 - International Work-conference on Time Series, Volume 2, Granada, Junho 2014.

Liu, J.; Xie, Z. (2014) BMA Probabilistic Quantitative Precipitation Forecasting over the Huaihe Basin Using TIGGE Multimodel *Ensemble* Forecasts, Monthly Weather Review, American Meteorological Society, 142, 2014.

Livino de Carvalho, A. R. (2001) Previsão de afluências diárias considerando horizonte de 1 a 7 dias através de uma abordagem estocástica – Modelo Previvazh. Tese de Mestrado em Engenharia Civil, COPPE/UFRJ, Rio de Janeiro, Brasil.

Lopes, J. E., Braga B.F.F., Conejo, J.L. (1982) A simplified hydrologic model in applied modeling in catchment hydrology, Water Resources Publication.

Maceira, M. E. P.; Penna, D. D. J.; Damázio, J. M. (2006) Geração de Cenários Sintéticos de Energia e Vazão para o Planejamento da Operação Energética. Cadernos do IME – Série Estatística, ISSN 1413-9022 / v. 21 p. 11 – 35, 2006.

Madigan, D.; Raftery, A. E. (1994) Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window. Journal of the American Statistical Association 89:1535–1546, 1994.

Maity, R.; Kumar, D. N. (2008) Basin-scale stream-flow forecasting using the information of large-scale atmospheric circulation phenomena. Hydrological Processes 22, páginas 643–650, 2008.

Meller, A.; Collischonn, W.; Fan, F.; Buarque, D. C.; Paiva, R. C. D.; Dias, P. L. S.; Moreira, D. S. (2014) Previsão de Cheias por Conjunto em Curto Prazo. Revista Brasileira de Recursos Hídricos, v. 19, n. 3, 2014.

Montgomery, J.; Nyhan, B. (2010) Bayesian Model Averaging: Theoretical developments and practical applications. Forthcoming, Political Analysis, março, 2010.

Noakes, D. J.; Mcleod, A. I.; Hipel, K. W. (1985) Forecasting Monthly Riverflow Time Series. International Journal of Forecasting 1, páginas 179-190, 1985.

Oliveira, G. C.; Kelman, J.; Pereira, M. V. F.; Stedinger, J. R. (1998) Synthetic Streamflow Generation Model Verification and Validation. Water Resources Research, vol. 24, nº 5, páginas 781-785, May 1988

ONS, Alteração do critério de escolha do PREVIVAZ com a utilização de técnicas de "Data Mining" para a bacia do Rio Uruguai, NT 133/2007. (2007a)

ONS, Novo modelo de previsão de vazões com informação de precipitação para o trecho incremental de Itaipu, NT 173/2007. (2007b)

ONS, Metodologia para a previsão de vazões uma semana à frente na bacia do alto/médio Rio Grande, Revisão 1 - Maio 2009. NT 139/2008. (2008)

ONS, RE 3/119/2011 Relatório Anual de Avaliação das Previsões de Vazões de 2010. (2011)

Paiva, L.F.G.; Acioli, G.C.L. (2007) Previsão de vazões com o modelo CPINS – Cálculo e Previsão de Vazões Naturais e Incrementais a Sobradinho. Seminário de Previsão de Vazões do ONS, 2007.

Piechota, T. C.; Chiew, F. H. S.; Dracup, J. A.; McMahon, T. A. (2001), Development of exceedance probability streamflow forecast, Journal of Hydrology Eng., 6(1), páginas 20–28, 2001.

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raftery, A. E. (1995) Bayesian Model Selection in Social Research. Sociological Methodology 25:111–163, 1995.

Raftery, A.E.; Gneiting, T.; Balabdaoui, F.; Polakowski, M. (2005) Using Bayesian Model Averaging to Calibrate Forecast *Ensembles*. American Meteorological Society, ed. 133, páginas 1155-1174. May, 2005.

Raftery, A.; Hoeting, J.; Volinsky, C.; Painter, I.; Yeung, K. Y. (2014) Package for Bayesian model averaging for linear models,generalizable linear models and survival models (Cox regression). version 3.17.1, 2014.
URL: http://www.r-project.org,http://www.research.att.com/~volinsky/bma.html

Reynolds, R.W.; Smith, T.M. (1994) Improved global sea surface temperature analyses using optimum interpolation. Journal of Climate 7, 929–948, 1994.

Robert, C. P.; Casella, G. (2004) *Monte Carlo statistical methods*. Springer, Nova York, 2004.

Salas, J.D.; Delleur, J.W.; Yejevich, V.; Lane, W.L. (1980) Applied modelling of hydrological time series. Water Resources Press Publications, Littleton, Colorado, E.U.A., 1980.

Salas, J.D. (1993) Analysis and Modeling of Hydrologic Time Series, Chapter 19 (72 p.) in The McGraw Hill Handbook of Hydrology, D.R. Maidment, Editor.

Sharma, A. (2000a) Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: 1. A strategy for system predictor identification, Journal of Hydrology, 239, páginas 232– 239, 2000.

Sharma, A. (2000b) Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: 3. A nonparametric probabilistic forecast model, Journal of Hydrology, 239, páginas 249–258, 2000.

Souza Filho, F. A.; Lall, U. (2003) Seasonal to interannual *ensemble* streamflow forecastsfor Ceara, Brazil: Applications of a multivariate, semiparametric algorithm. Water Resources Research, vol. 39, nº 11, 1307. 2003.

Srinivas, V.V.; Srinivasan, K. (2005) Hybrid moving block bootstrap for stochastic simulation of multi-site streamflows. Journal of Hydrology 302, páginas 307–330, 2005.

Stedinger, J. R.; Taylor, M. R. (1982) A Representation of Spatial Cross Correlations in Large Stochastic Seasonal Streamflow Models. Water Resources Research, vol. 18, nº4, August 1982.

Sveinsson, O.G.B.; Salas, J.D. (2003) Modeling of dynamics of long-term variability of hydroclimatic processes. Journal of Hydrometeorology 4(3):489–505, 2003.

Tesfaye, Y.G.; Meerschaert, M.M.; Anderson P.L. (2006) Identification of periodic autoregressive moving average models and their application to the modeling of river flows. Water Resources Research, vol. 42, 2006

Tibshirani (1996) Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 58, No. 1, páginas 267-288. 1996.

Tippett, M. K.; Kleeman, R.; Tang, Y. (2004) Measuring the potential utility of seasonal climate predictions. Geophysical Research Letters, vol. 31, 2004

Tootle, G. A.; Piechota, T. C.; Gutiérrez, F. (2008) The relationships between Pacific and Atlantic Ocean sea surface temperatures and Colombian streamflow variability. Journal of Hydrology, p. 268– 276, 2008.

Uvo, C.B., Graham, N.E. (1998) Seasonal runoff forecast for northern South America: a statistical model. Water Resources Research 34 (12), p. 3515–3524, 1998.

Van der Kooij, A.J.; Meulman, J.J. (2006) Regularization with Ridge penalties, the Lasso, and the Elastic Net for Regression with Optimal Scaling Transformations. https://openaccess.leidenuniv.nl/bitstream/handle/1887/12096/04.pdf?sequence=18. Acesso em 05/12/2014.

Voisin, N.; Hamlet, A. F.; Graham, J. P.; Pierce, D. W.;Barnett, T. P.; Lettenmaier, D. P. (2006) The Role of Climate Forecasts in Western U.S. Power Planning. Journal of Applied Meteorology and Climatology, volume 45, Maio 2006.

Vrugt, J. A.; Diks, C. G. H.; Clark, M. P. (2008) *Ensemble* Bayesian model averaging using Markov Chain Monte Carlo sampling. Environ Fluid Mech, DOI 10.1007/s10652-008-9106-3, 2008.

Wang, Q. J.; Schepen, A.; Robertson, D. E. (2012) Merging Seasonal Rainfall Forecasts from Multiple Statistical Models through Bayesian Model Averaging. Journal of Climate, 25, 2012.

Westra, S.; Brown, C.; Lall, U.; Sharma, A. (2007) Modeling multivariable hydrological series: Principal component analysis or independent component analysis? Water Resources Research, vol. 43, 2007.

Wintle, B.A.; McCarthy, M.A.; Volinsky, C.T.; Kavanagh, R.P. (2003) The Use of Bayesian Model Averaging to Better Represent Uncertainty in Ecological Models. Conservation Biology 17:1579–1590, 2003.

Yee, Y. K.; Bumgarner, R. E.; Raftery, A. E. (2005) Bayesian model averaging: development of an improved multi-class, gene selection and classification tool for microarray data. Bioinformatics 21:2394–2402, 2005.

Feldkircher, M.; S. Zeugner (2009) Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging; IMF Working Paper 09-202, 2009.

APÊNDICES

O presente tópico foi dividido em cinco subitens: o primeiro apêndice traz um exemplo didádico da aplicação da técnica de ponderação bayesiana em um caso hipotético com dois modelos, extraído do artigo de Gibbons *et al*. (2008); o segundo, apresenta, de forma geral, a formulação do algoritmo *expectation-maximization* (EM), desenvolvido por Dempster *et al.* (1977), acompanhado de um caso ilustrativo, retirado do artigo de Do e Batzoglou (2008); o Apêndice C traz os resultados anuais dos modelos utilizados pelo ONS que foram utilizados para a confecção dos gráficos das Figuras 6.16 a 6.18 desta dissertação; no quarto apêndice constam as figuras com os resultados dos três principais indicadores utilizados na análise do desempenho das previsões para cada um dos vinte e oito reservatórios estudados nesta dissertação. Os indicadores ilustrados nas Figuras D.1 a D.28 são: os pesos bayesianos para os modelos semanais e mensais; o indicador distância multicritério (DM) das previsões mensais do modelo PARX e o indicador DM para as previsões semanais antes e após as ponderações BMA. Por fim, o último dos apêndices apresenta um comparativo do desempenho dos modelos semanais utilizados pelo ONS e os resultados após a utilização da técnica BMA no período compreendido entre jan-2009 e set-2014, por meio dos quatro indicadores de qualidade médios adotados nesta dissertação. Esses dados, que constam das Tabelas E.1 a E.4, serviram de base para a confecção das Figuras 6.26 a 6.29 apresentadas anteriormente.

APÊNDICE A - EXEMPLO SIMPLIFICADO DE APLICAÇÃO DA TÉCNICA BMA (EXTRAÍDO DE GIBBONS *ET AL.*, 2008)

Considere um par de modelos (M_1 e M_2) com apenas um parâmetro cada, de modo a ilustrar, de forma simplificada, uma aplicação da técnica BMA. Considere que esses modelos possuem o mesmo conjunto de valores (n) de dados (D) gerados por um modelo "verdadeiro", M_0. A motivação do exemplo é identificar a forma mais adequada de geração de dados com a mesma distribuição de D, utilizando M_1 e M_2 de forma individual ou em alguma combinação dos dois (ou seja, faremos uma suposição a mais próxima do real de que conhecemos D, mas não conhecemos M_0). Ambos os modelos seguem distribuições normais N [μ, σ^2] com parâmetros $M_0 \sim N(0.0, 1.0)$. No caso de M_1 e M_2, as médias μ são desconhecidas (deverão ser estimadas a partir de D) enquanto que os desvios (σ) são conhecidos ($\sigma_1 = 1.0$ e $\sigma_2 = 1.1$). O exemplo foi disposto de modo que o modelo M_1 pode ser equivalente a M_0 (se $\mu_1 = 1.0$), mas M_2 não pode, independentemente do valor do μ_2. Se estiverem disponíveis dados suficientes, espera-se identificar M_1 como o modelo mais adequado, mas se os dados forem poucos, será mais difícil distinguir M_1 de M_2.

Para aplicar inferência bayesiana e BMA neste problema, deve-se especificar uma *priori* para μ. Para os fins deste exemplo, adotou-se uma distribuição normal "não informativa" com $\mu = 0{,}0$ e $\sigma = 1{,}0 \times 10^3$, o que permite calcular diretamente as verossimilhanças marginais, e, portanto, as probabilidades a *posteriori* dos modelos (PMPs – *Posterior Model Probabilities*). Na maioria das situações práticas, não existirá uma forma analítica, assim deverá ser aplicada alguma técnica de amostragem (ex: MCMC). Também devemos especificar uma probabilidade a *priori* para cada modelo. Por simplicidade, assumimos o valor 0.5, ou seja, nenhum modelo se sobrepõe ao outro, a priori.

Para i= 1, 2, considere que o Modelo i: $x_1, \dots, x_n$ é independente e identicamente distribuído (i.i.d.) com distribuição N(μ, σ_i^2) e que a distribuição a *priori* de μ é N(a_i, b_i^2). Os valores de σ_1, σ_2, a_1, a_2, b_1 e b_2 são assumidos como conhecidos e constantes. Por simplicidade, assumimos que $a_1 = a_2 = 0$ e $b_1 = b_2 = \gamma$. Escrevendo:

$$S = \sum_{i=1}^{n} x_i^2$$

e (A.1)

$$T = \sum_{i=1}^{n} x_i$$

Fazendo p_1 e p_2 denotarem as probabilidades a *priori* para o Modelos 1 e 2, respectivamente, então $p_1 + p_2 = 1$. A distribuição a *posteriori* de μ para o Modelo i, considerando a distribuição normal, é dada por:

$$p_i(\mu|D, \sigma_i) = N\left(\frac{\frac{T}{\sigma_i^2}}{\left(\frac{1}{\gamma^2} + \frac{n}{\sigma_i^2}\right)}, \frac{1}{\left(\frac{1}{\gamma^2} + \frac{n}{\sigma_i^2}\right)}\right) \quad (A.2)$$

E a verossimilhança marginal para o Modelo i é dada por:

$$f_i(x|\gamma) = \left(\frac{1}{2\pi\sigma_i^2}\right)^{n/2} \left(1 + \frac{n\gamma^2}{\sigma_i^2}\right)^{-1/2} \times exp\left[-\frac{1}{2}\frac{\sigma_i^{-4}(S - n^{-1}T^2) + n^{-1}\sigma_i^{-2}\gamma^{-2}S}{\sigma_i^{-2} + n^{-1}\gamma^{-2}}\right] \quad (A.3)$$

Do teorema de Bayes, as probabilidades a *posteriori* do modelo w_1 e w_2 são dadas por:

$$w_i = \frac{p_i f_i(x|\gamma)}{p_1 f_1(x|\gamma) + p_2 f_2(x|\gamma)}, \quad para\ i = 1, 2 \quad (A.4)$$

Pela lei dos grandes números, se o tamanho da amostra **n** aumenta, a variância da amostra $n^{-1}S - n^{-2}T^2$ converge para a variância da população σ^2. Sob o Modelo 1, $\sigma^2 = \sigma_1^2$ e sob o Modelo 2, $\sigma^2 = \sigma_2^2$. Além disso, pode ser verificado que, sob o Modelo 1 com probabilidade 1, a relação de $f_1(x_j \mid \gamma) / f_2(x_j \mid \gamma)$ aumenta até ao infinito quando **n** tende ao infinito. Consequentemente, sob o Modelo 1 com probabilidade 1, w_1 vai para 1 e w_2 vai para 0 quando **n** vai para o infinito, ocorrendo o inverso, sob o Modelo 2.

O exemplo da Figura A.1 ilustra três tamanhos de conjuntos de dados (n = 5, 10 e 100). Para n = 100, a verossimilhança marginal de M_1 é 5.77×10^{66} e para M_2 é $2{,}31 \times 10^{66}$. Assim, o valor de PMP para M_1 é 5,77 / (5,77 + 2,31) = 0,71 e para M_2 é 2,31 / (5,77 + 2,31) = 0,29. Na Figura A.1, vemos que a *posteriori*, a verossimilhança e a verossimilhança marginal mudam à medida que **n** muda. Quanto menores os valores de **n**,

existe menos informação nos dados e os valores de PMP aproximam-se dos valores a priori. Como demonstrado acima, quando **n** tende ao infinito, os valores de PMP para M_1 tendem para 1. Nota-se que, embora a assunção da *priori* para μ seja "difusa" (ao longo de um amplo espectro), a distribuição a *posteriori* não é, mesmo com baixos valores de **n**. Este resultado ocorre porque a verossimilhança recebe pouco suporte através dos dados de μ. Uma vez que os valores de PMP foram estimados, os modelos podem ser utilizados para fazer previsões conjuntas de D. No caso de n = 100, as previsões conjuntas de D seriam compostas de 71% das amostras de M_1 e 29% das amostras de M_2.

A ideia estende-se a qualquer número de modelos. Por exemplo, se um terceiro modelo M_3 for adicionado com uma probabilidade marginal de $3{,}00 \times 10^{66}$, os valores PMPs seriam atualizados, para M_3 o cálculo seria de 3,00 / (5,77 + 2,31+ 3,00) = 0,27. Observe que, bem como alterar o valor de **n**, o valor de PMP não é absoluto, ele possui relação com os outros modelos da série.

Na Figura A.1, os números circulados ao lado das setas indicam as etapas do processo: a Etapa 1 é a estimativa da verossimilhança a partir dos modelos e dados; a Etapa 2 é o cálculo da verossimilhança marginal por integração da verossimilhança e da *priori*; a Etapa 3 é a multiplicação da *priori* e da verossimilhança para formar o numerador na equação de Bayes; a Etapa 4 é a normalização do numerador pelo denominador, para estimar a distribuição a *posteriori*. Os maiores gráficos representam a condição em que n = 100, os gráficos internos ilustram o efeito da alteração do valor de **n**, sendo que o gráfico superior é para n = 10 e o inferior para n = 5. Note que a escala de abscissas para a densidade a *priori* difere dos outros gráficos de densidade. Observe também que a *priori* é elaborada a partir de uma distribuição normal, mas não aparece em forma de sino porque foi plotada em um intervalo estreito em comparação com a variância.

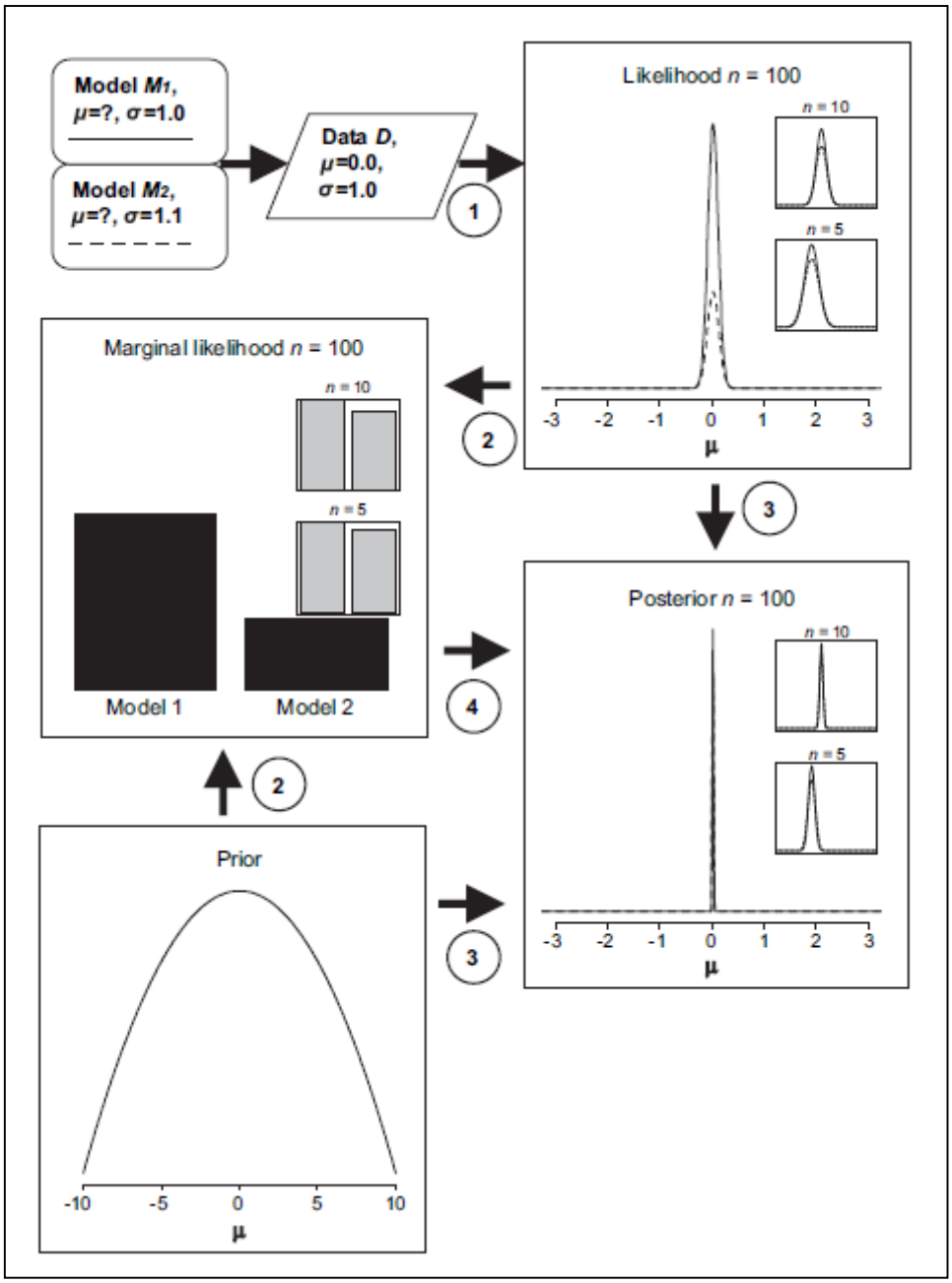

Figura A.1 - Um exemplo de inferência bayesiana utilizando dois modelos de distribuição normal M1 e M2 ajustados a um conjunto de dados gerados normalmente, D, de tamanho **n**. (Fonte: Gibbons *et al*., 2008)

APÊNDICE B - O ALGORITMO *EXPECTATION-MAXIMIZATION*

O algoritmo *expectation-maximization* (EM) é um procedimento iterativo de cálculo da máxima verossimilhança quando apenas um subconjunto do conjunto completo de dados está disponível. Dempster *et al.* (1977) demonstraram a ampla aplicabilidade do algoritmo EM e são em grande parte responsáveis por popularizar este método na estatística. Maiores detalhes do algoritmo EM podem ser encontrados em Hastie *et al.* (2008) e Vrugt *et al.* (2008).

Para auxiliar na compreensão do algoritmo, considere a experiência de lançamento de um par de moedas A e B, extraída de Do e Batzoglou (2008). Ambas as moedas possuem vieses desconhecidos: θ_A, para a moeda A, e θ_B, para a moeda B (isto é, em qualquer lançamento, a probabilidade de o resultado ser cara para a moeda A é θ_A e a probabilidade de ser coroa é 1- θ_A, similarmente para a moeda B). O objetivo é estimar $\theta = (\theta_A, \theta_B)$ repetindo o procedimento a seguir cinco vezes: escolhe-se aleatoriamente uma das duas moedas (com igual probabilidade) e realizam-se dez lançamentos independentes com a moeda selecionada. Assim, todo o processo envolve um total de 50 lançamentos, conforme a Figura B.1-a. Para cada conjunto de dez lançamentos, o procedimento de máxima verossimilhança acumula a contagem de caras e coroas para as moedas A e B separadamente. Estas contagens são então utilizadas para estimar o viés de cada moeda.

Durante a experiência, suponha que são mantidos o controle de dois vetores $x = (x_1, x_2, ..., x_5)$ e $z = (z_1, z_2, ..., z_5)$, em que $x_i \in \{0,1, ..., 10\}$ é o número de caras observadas durante o i-ésimo lançamento, e $z_i \in \{A, B\}$ é a identidade da moeda utilizada durante o i-ésimo lançamento. A estimação de parâmetros neste cenário é conhecida como o "caso completo de dados", em que os valores de todas as variáveis aleatórias relevantes do nosso modelo são conhecidos (que é o resultado de cada lançamento de moeda e do tipo de moeda usada para cada jogada). Aqui, uma maneira simples de estimar θ_A e θ_B é calcular as proporções observadas de caras para cada moeda:

$$\hat{\theta}_A = \frac{número\ de\ caras\ utilizando\ a\ moeda\ A}{total\ de\ lançamentos\ utilizando\ a\ moeda\ A}$$

e (B.1)

$$\hat{\theta}_B = \frac{número\ de\ caras\ utilizando\ a\ moeda\ B}{total\ de\ lançamentos\ utilizando\ a\ moeda\ B}$$

Esta suposição é conhecida na literatura como a estimativa da máxima verossimilhança (o método de máxima verossimilhança avalia a qualidade de um modelo estatístico baseada na probabilidade que o modelo atribui aos dados observados). Se logP (x, z; θ) é o logaritmo da probabilidade conjunta (log-verossimilhança) de se obter qualquer vetor específico de observações de caras x e de moedas selecionadas z, então as equações em (B.1) são a solução para os parâmetros $\hat{\theta} = (\hat{\theta}_A, \hat{\theta}_B)$ que maximizem logP (x, z; θ).

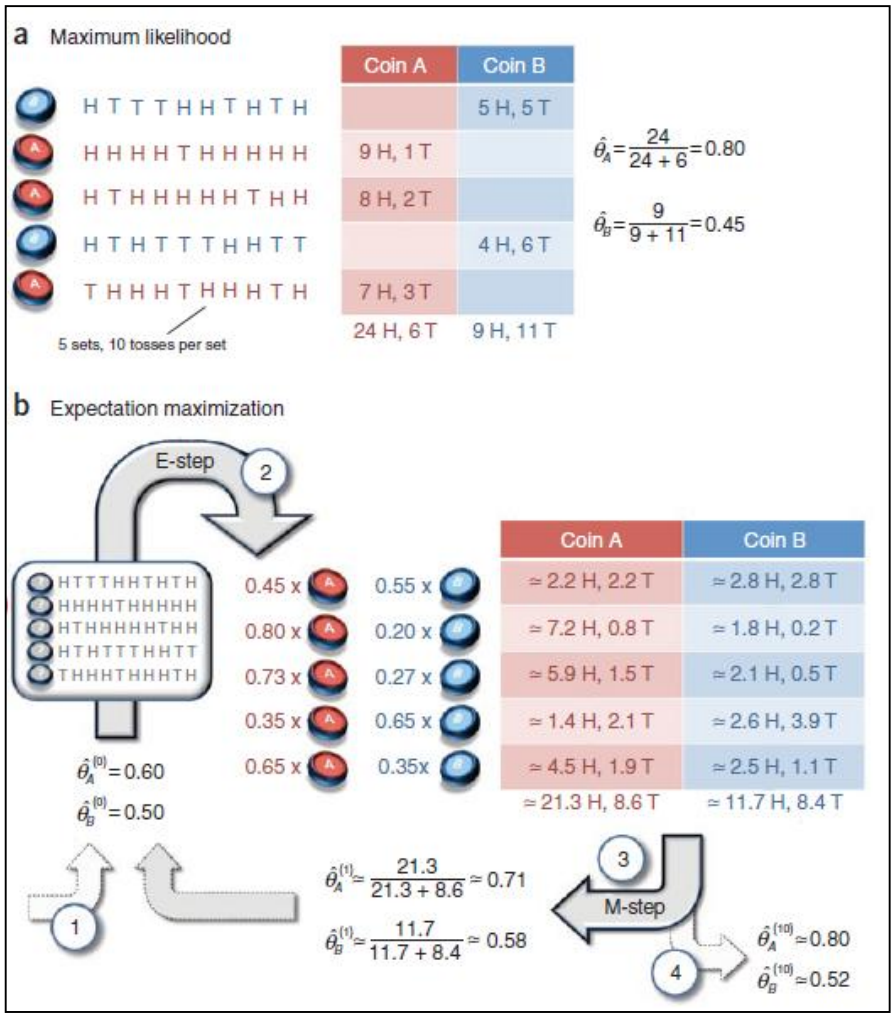

Figura B.1 - Estimativa de parâmetros para dados completos e incompletos. (a) estimativa de máxima verossimilhança. (b) O passo a passo do algoritmo EM. (fonte: Do e Batzoglou, 2008).

Considere então uma variante do problema de estimação de parâmetros acima, em que nos é fornecida a contagem de caras obtidas x, mas não as informações z sobre quais moedas foram selecionadas para cada conjunto de lançamentos. Neste caso, as variáveis z

são chamadas de variáveis ocultas ou fatores latentes. A estimação de parâmetros neste novo cenário é conhecida como o "caso incompleto de dados". Desta vez, estimar as proporções de caras para cada moeda não é mais possível, porque não sabemos a moeda usada para cada conjunto de lançamentos. No entanto, se tivéssemos alguma forma de completar os dados (no nosso caso, adivinhar corretamente que moeda foi utilizada em cada uma das cinco sequencias de lançamentos), então poderíamos reduzir a estimativa de parâmetros para este problema com dados incompletos por meio da estimativa de máxima verossimilhança do caso com os dados completos.

Um esquema iterativo para a obtenção dos preenchimentos de falhas poderia funcionar da seguinte forma: a partir de alguns parâmetros iniciais, $\hat{\theta}^{(t)} = (\hat{\theta}_A^{(t)}, \hat{\theta}_B^{(t)})$, determinar, para cada um dos cinco conjuntos, em qual das moedas A ou B é mais provável gerar-se os lançamentos observados (usando-se os parâmetros atuais estimados). Em seguida, assumindo esses preenchimentos como adequados (ou seja, a hipótese de que a seleção das moedas foi adequada), aplica-se o procedimento de estimação de máxima verossimilhança para obter $\hat{\theta}^{(t+1)}$. Finalmente, repetem-se estes dois passos até a convergência. À medida que o modelo estimado melhora, a qualidade dos preenchimentos resultantes também irá melhorar.

O algoritmo expectation-maximization (EM) é um refinamento dessa ideia básica. Ao invés de adotar um único preenchimento mais provável para a seleção das moedas em cada iteração, o algoritmo EM calcula as probabilidades para cada preenchimento possível dos dados, usando os parâmetros atuais $\hat{\theta}^{(t)}$. Estas probabilidades são usadas para criar um conjunto ponderado de treinamento, consistindo de todos os possíveis preenchimentos dos dados. Finalmente, uma versão modificada da estimativa da máxima verossimilhança que lida com exemplos de treinamento ponderados fornecerá novas estimativas de parâmetros, $\hat{\theta}^{(t+1)}$. Usando exemplos ponderados de treinamento em vez de escolher o único melhor preenchimento, o algoritmo EM explica a confiança do modelo em cada preenchimento dos dados (Figura B.1-b).

Em resumo, o algoritmo EM alterna entre dois passos: o primeiro passo consiste na estimação de uma distribuição de probabilidade para o preenchimento de dados faltantes, dado o modelo atual (conhecido como o passo "expectativa", ou o passo-E); o segundo implica em reestimar os parâmetros do modelo utilizando esses preenchimentos (conhecido como o passo "maximização", ou passo-M). O nome 'passo-E' vem do fato de

ser geralmente um passo não necessário para compor a distribuição de probabilidade para o preenchimento de forma explícita, precisando computar apenas as estatísticas suficientes "esperadas" acerca desses preenchimentos. Da mesma forma, o ‘passo-M’ é assim denominado em função de a reestimação do modelo poder ser pensada como a “maximização” da log-verossimilhança esperada dos dados.

Na Figura B.1-b, são elencados os 4 passos do algoritmo EM para o exemplo: 1) O algoritmo começa com uma estimativa inicial dos parâmetros. 2) No passo-E, uma distribuição de probabilidade sobre os possíveis preenchimentos é calculada usando os parâmetros atuais. As contagens mostradas na tabela são os números esperados de caras e coroas de acordo com esta distribuição. 3) No passo-M, novos parâmetros são determinados usando as conclusões atuais. 4) Depois de várias repetições dos passos E e M, o algoritmo converge.

APÊNDICE C - DESEMPENHO DAS PREVISÕES DOS MODELOS UTILIZADOS PELO ONS EM RELAÇÃO ÀS VAZÕES OBSERVADAS (DISTÂNCIA MULTICRITÉRIO ANUAL ENTRE 2009 E 2013).

		1 semana de antecedência					
Rio	**Reservatório**	**2009**	**2010**	**2011**	**2012**	**2013**	**Média**
Grande	Furnas	0,45	0,47	0,23	0,29	0,34	**0,36**
	M. Moraes	0,44	0,40	0,22	0,29	0,43	**0,36**
	Marimbondo	0,27	0,17	0,30	0,31	0,53	**0,32**
	Água Vermelha	0,24	0,16	0,17	0,28	0,53	**0,28**
Paranaíba	Emborcação	0,40	0,65	0,62	0,23	0,35	**0,45**
	Nova Ponte	0,56	1,31	0,34	0,29	0,75	**0,65**
	Itumbiara	0,35	0,43	0,40	0,24	0,45	**0,37**
	São Simão	0,14	0,17	0,20	0,15	0,29	**0,19**
Paranapanema	Jurumirim	0,15	0,15	0,20	0,15	0,20	**0,17**
	Chavantes	0,46	0,32	0,80	1,18	2,05	**0,96**
	Capivara	0,75	0,57	1,11	2,08	1,90	**1,28**
Iguaçu	Foz do Areia	0,58	0,47	1,08	1,41	0,69	**0,85**
	Salto Santiago	0,51	0,57	0,39	0,69	0,45	**0,52**
Uruguai	Itá	0,51	0,65	0,41	0,77	0,53	**0,57**
	Passo Fundo	0,76	0,79	0,78	0,76	0,75	**0,77**
	Machadinho	0,50	1,16	0,85	1,94	1,68	**1,23**
Jacuí	Passo Real	0,57	1,97	1,81	1,26	1,44	**1,41**
São Francisco	Três Marias	0,35	0,32	0,22	0,35	0,28	**0,30**
	Sobradinho	0,07	0,05	0,05	0,05	0,05	**0,05**
	Itaparica	0,13	0,07	0,06	0,05	0,07	**0,08**
Tietê	Barra Bonita	0,74	0,80	0,75	0,76	0,66	**0,74**
	Promissão	0,45	0,31	0,84	0,93	0,97	**0,70**
	Três Irmãos	0,43	0,28	0,47	0,75	0,86	**0,56**
Paraná	Ilha Solteira	0,38	0,32	0,62	0,71	0,87	**0,58**
	Jupiá	0,14	0,14	0,24	0,17	0,23	**0,18**
	Itaipu	0,15	0,12	0,26	0,20	0,29	**0,20**
Tocantins	Serra da Mesa	0,44	0,36	0,48	0,30	0,51	**0,42**
	Tucuruí	0,09	0,12	0,16	0,13	0,11	**0,12**

Tabela C.1 – Resultados do indicador DM com uma semana de antecedência. Fonte: elaboração própria, adaptado de www.ons.org.br.

		2 semanas de antecedência					
Rio	**Reservatório**	**2009**	**2010**	**2011**	**2012**	**2013**	**Média**
Grande	Furnas	0,50	0,60	0,53	0,75	1,09	**0,69**
	M. Moraes	0,48	0,56	0,53	0,68	1,04	**0,66**
	Marimbondo	0,39	0,37	0,39	0,56	0,59	**0,46**
	Água Vermelha	0,39	0,33	0,56	0,51	0,61	**0,48**
Paranaíba	Emborcação	0,52	0,71	0,69	0,35	0,67	**0,59**
	Nova Ponte	0,66	0,77	0,44	0,50	0,99	**0,67**
	Itumbiara	0,48	0,67	0,53	0,38	0,55	**0,52**
	São Simão	0,30	0,37	0,47	0,29	0,47	**0,38**
Paranapanema	Jurumirim	0,27	0,33	0,35	0,27	0,33	**0,31**
	Chavantes	0,60	0,60	1,02	5,41	4,54	**2,43**
	Capivara	0,76	0,87	1,15	4,94	2,58	**2,06**
Iguaçu	Foz do Areia	0,92	0,41	1,26	1,85	1,24	**1,14**
	Salto Santiago	0,80	0,80	0,68	1,20	1,00	**0,90**
Uruguai	Itá	0,79	0,96	0,58	1,20	1,06	**0,92**
	Passo Fundo	0,82	0,93	0,98	1,30	1,10	**1,03**
	Machadinho	0,72	1,52	1,25	1,60	2,34	**1,49**
Jacuí	Passo Real	0,80	1,37	1,47	1,41	1,45	**1,30**
São Francisco	Três Marias	0,46	0,69	0,52	0,56	0,62	**0,57**
	Sobradinho	0,13	0,12	0,10	0,12	0,11	**0,12**
	Itaparica	0,16	0,09	0,09	0,12	0,11	**0,11**
Tietê	Barra Bonita	0,87	0,85	0,94	1,23	1,06	**0,99**
	Promissão	0,56	0,41	0,82	0,92	1,06	**0,75**
	Três Irmãos	0,58	0,47	0,67	0,73	0,89	**0,67**
Paraná	Ilha Solteira	0,58	0,53	0,60	0,83	1,04	**0,72**
	Jupiá	0,24	0,29	0,38	0,27	0,31	**0,30**
	Itaipu	0,34	0,16	0,36	0,50	0,49	**0,37**
Tocantins	Serra da Mesa	0,59	0,46	0,39	0,45	0,53	**0,48**
	Tucuruí	0,17	0,15	0,19	0,20	0,15	**0,17**

Tabela C.2 – Resultados do indicador DM com duas semanas de antecedência. Fonte: elaboração própria, adaptado de www.ons.org.br.

		3 semana de antecedência					
Rio	**Reservatório**	**2009**	**2010**	**2011**	**2012**	**2013**	**Média**
Grande	Furnas	0,44	0,46	0,49	0,44	0,67	**0,50**
	M. Moraes	0,45	0,47	0,46	0,42	0,67	**0,49**
	Marimbondo	0,37	0,34	0,36	0,50	0,67	**0,45**
	Água Vermelha	0,37	0,30	0,35	0,42	0,63	**0,41**
Paranaíba	Emborcação	0,44	0,57	0,79	0,36	0,62	**0,56**
	Nova Ponte	0,37	1,30	0,39	0,40	1,30	**0,75**
	Itumbiara	0,29	0,67	0,57	0,39	0,48	**0,48**
	São Simão	0,34	0,41	0,48	0,29	0,46	**0,40**
Paranapanema	Jurumirim	0,33	0,31	0,39	0,33	0,35	**0,34**
	Chavantes	0,89	0,62	0,80	1,61	3,15	**1,41**
	Capivara	0,00	0,82	0,92	1,57	2,29	**1,12**
Iguaçu	Foz do Areia	0,00	0,63	0,88	1,14	1,34	**0,80**
	Salto Santiago	1,01	1,15	0,95	1,33	1,29	**1,15**
Uruguai	Itá	0,90	1,24	0,80	1,40	1,22	**1,11**
	Passo Fundo	1,11	1,39	1,08	1,58	1,01	**1,23**
	Machadinho	1,01	1,38	1,38	1,46	1,62	**1,37**
Jacuí	Passo Real	0,80	1,11	1,27	1,09	1,25	**1,10**
São Francisco	Três Marias	0,34	0,87	0,65	0,60	0,59	**0,61**
	Sobradinho	0,21	0,34	0,16	0,26	0,26	**0,25**
	Itaparica	0,21	0,27	0,13	0,23	0,22	**0,21**
Tietê	Barra Bonita	1,18	1,34	1,03	1,51	1,01	**1,21**
	Promissão	0,54	0,60	0,90	0,89	0,66	**0,72**
	Três Irmãos	0,58	0,47	0,65	0,92	0,70	**0,66**
Paraná	Ilha Solteira	0,53	0,44	0,66	0,93	0,77	**0,67**
	Jupiá	0,27	0,26	0,38	0,34	0,35	**0,32**
	Itaipu	0,48	0,20	0,52	0,69	0,67	**0,51**
Tocantins	Serra da Mesa	0,59	0,55	0,45	0,33	0,79	**0,54**
	Tucuruí	0,23	0,18	0,18	0,23	0,17	**0,20**

Tabela C.3 – Resultados do indicador DM com três semanas de antecedência. Fonte: elaboração própria, adaptado de www.ons.org.br.

APÊNDICE D – DESEMPENHO DAS SIMULAÇÕES POR RESERVATÓRIO

Bacia do rio Paranaíba

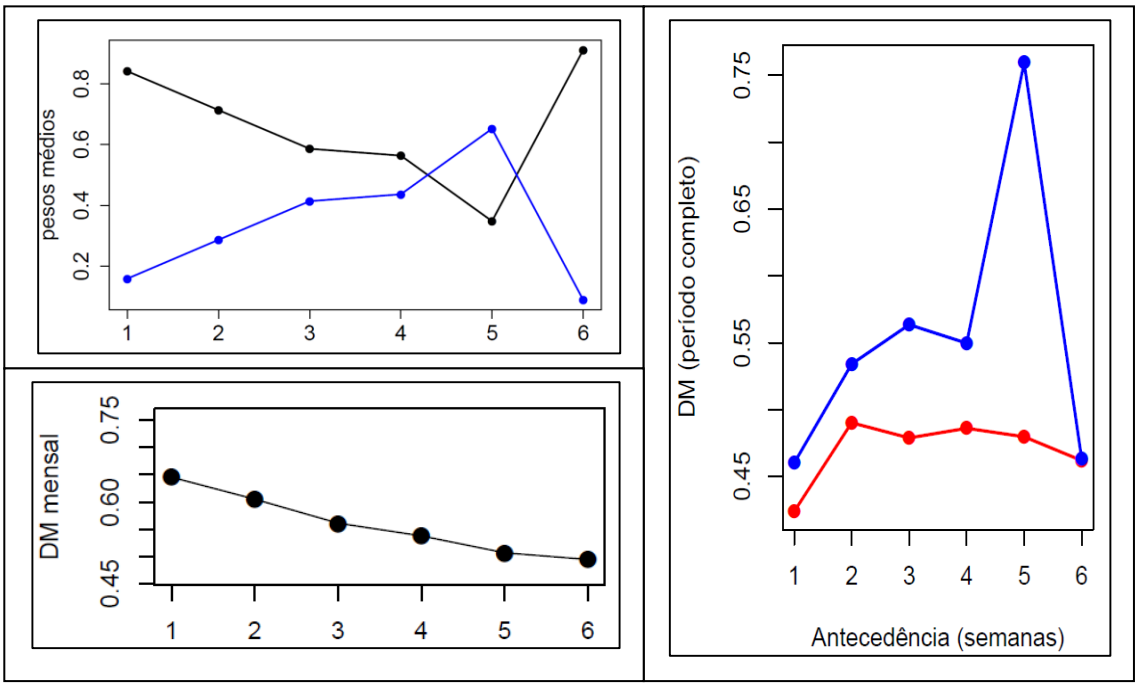

Figura D.1 - Resultados de uma a seis semanas de antecedência para a UHE Nova Ponte no período entre 2009 e 2014: o painel superior da esquerda mostra a ponderação dos pesos dos modelos semanais (curva na cor preta) e mensais (cor azul); o painel inferior da esquerda apresenta o indicador DM das previsões mensais (convertidos para a escala semanal). Por fim, o painel da direita traz o indicador DM para as previsões semanais realizadas pelo ONS (curva na cor azul) e para as previsões BMA (curva vermelha).

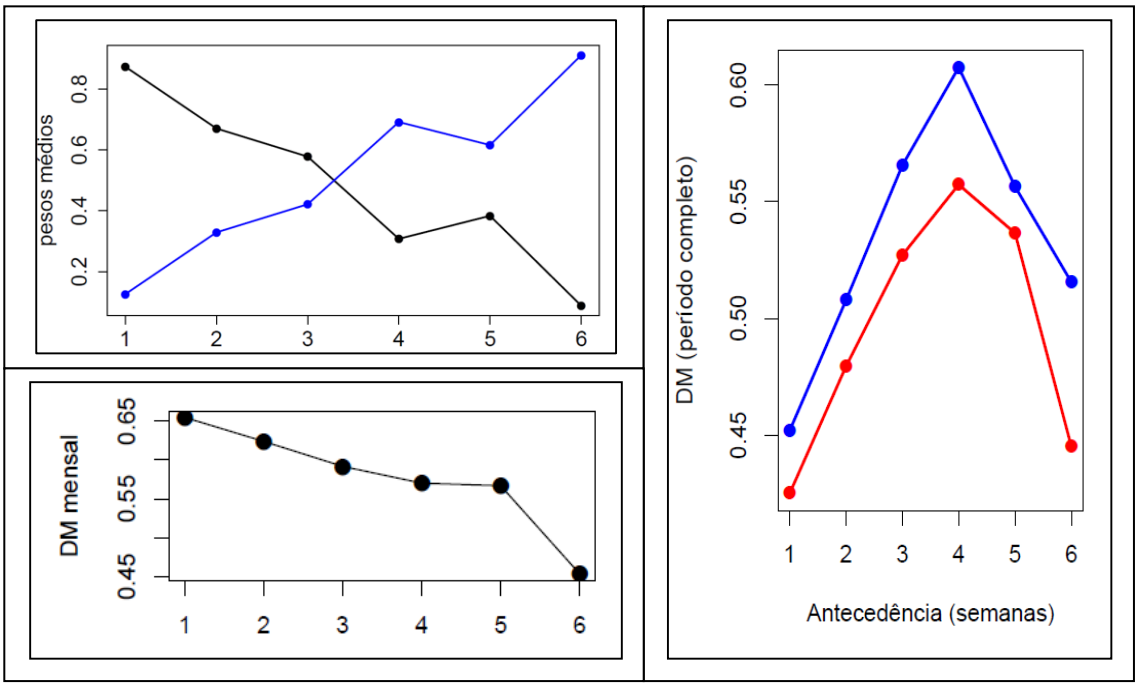

Figura D.2 - Assim como na Figura D.1, mas para a UHE Emborcação.

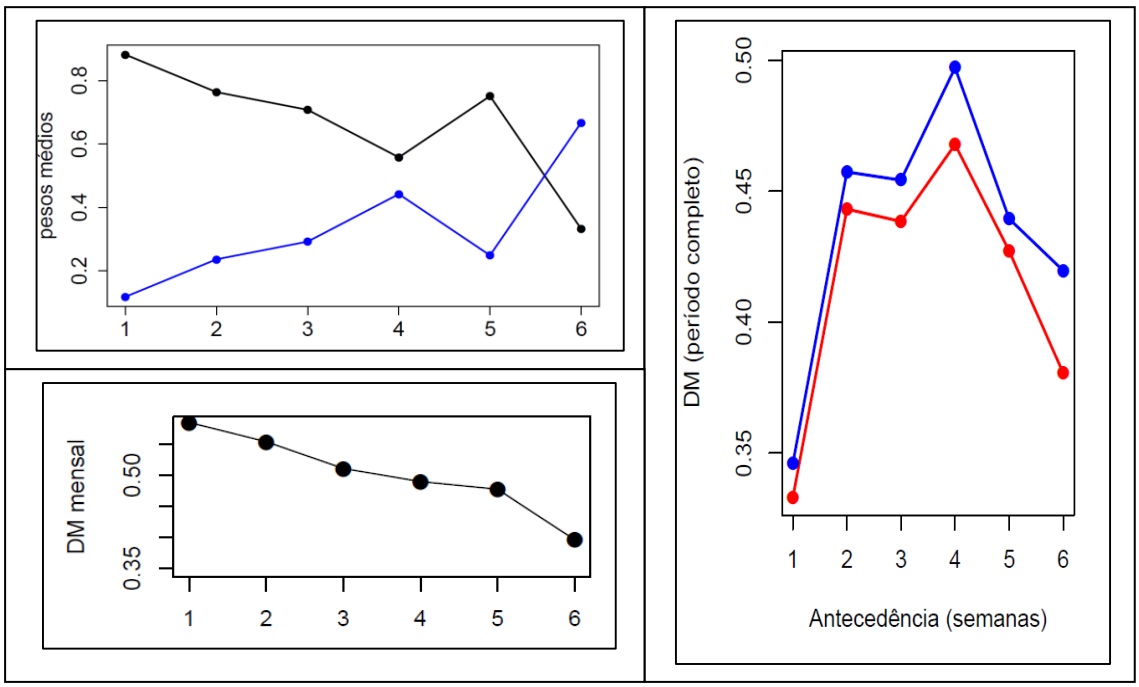

Figura D.3 - Assim como na Figura D.1, mas para a UHE Itumbiara.

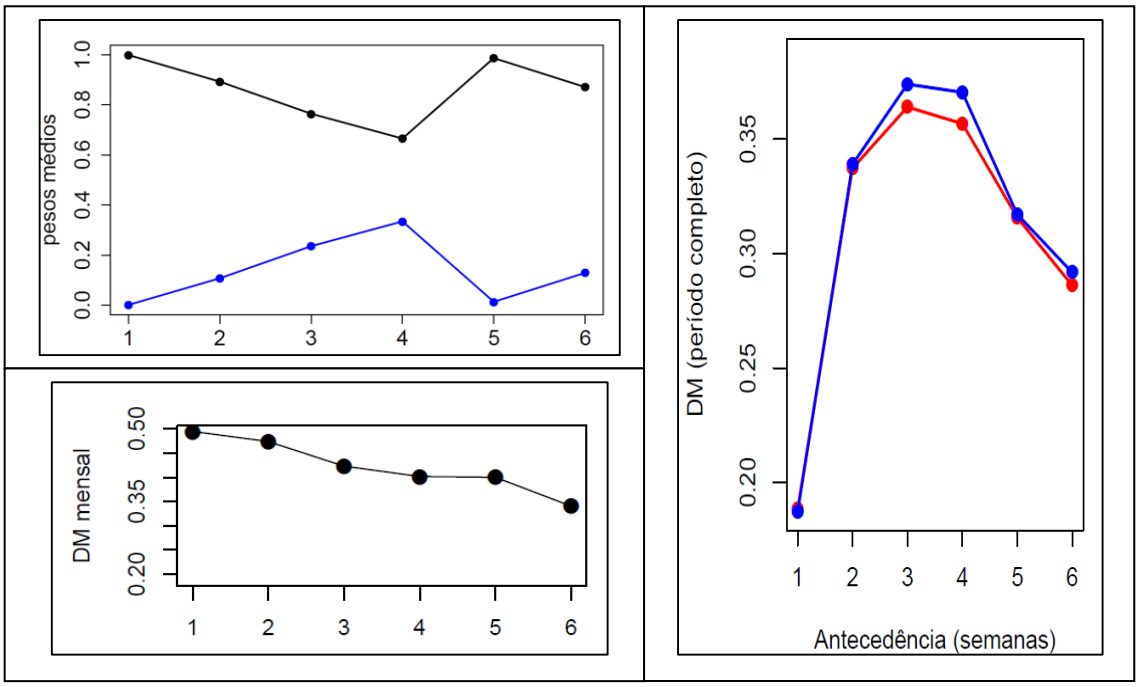

Figura D.4 - Assim como na Figura D.1, mas para a UHE São Simão.

Bacia do rio Grande

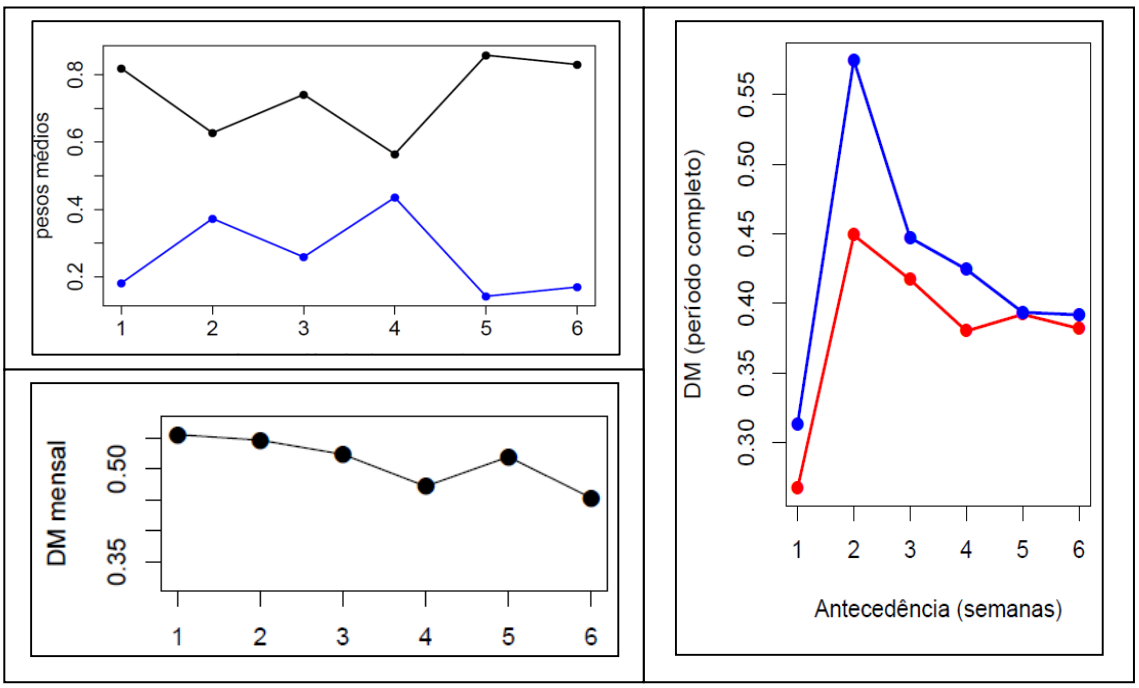

Figura D.5 - Assim como na Figura D.1, mas para a UHE Furnas.

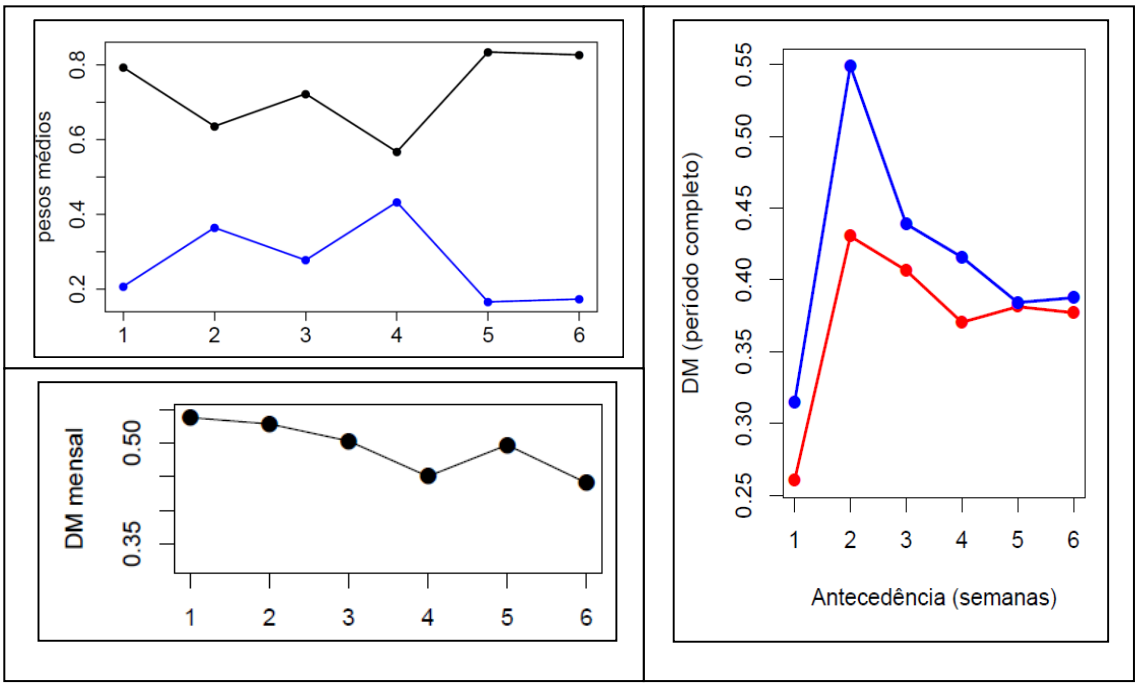

Figura D.6 - Assim como na Figura D.1, mas para a UHE Mascarenhas de Moraes.

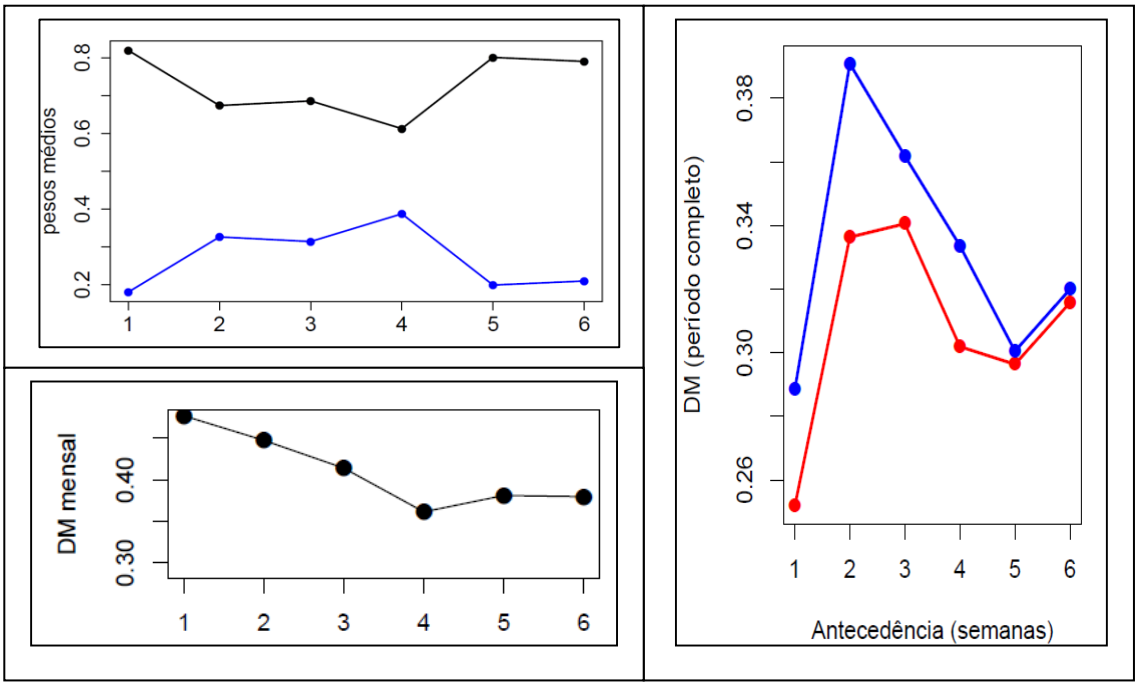

Figura D.7 - Assim como na Figura D.1, mas para a UHE Marimbondo.

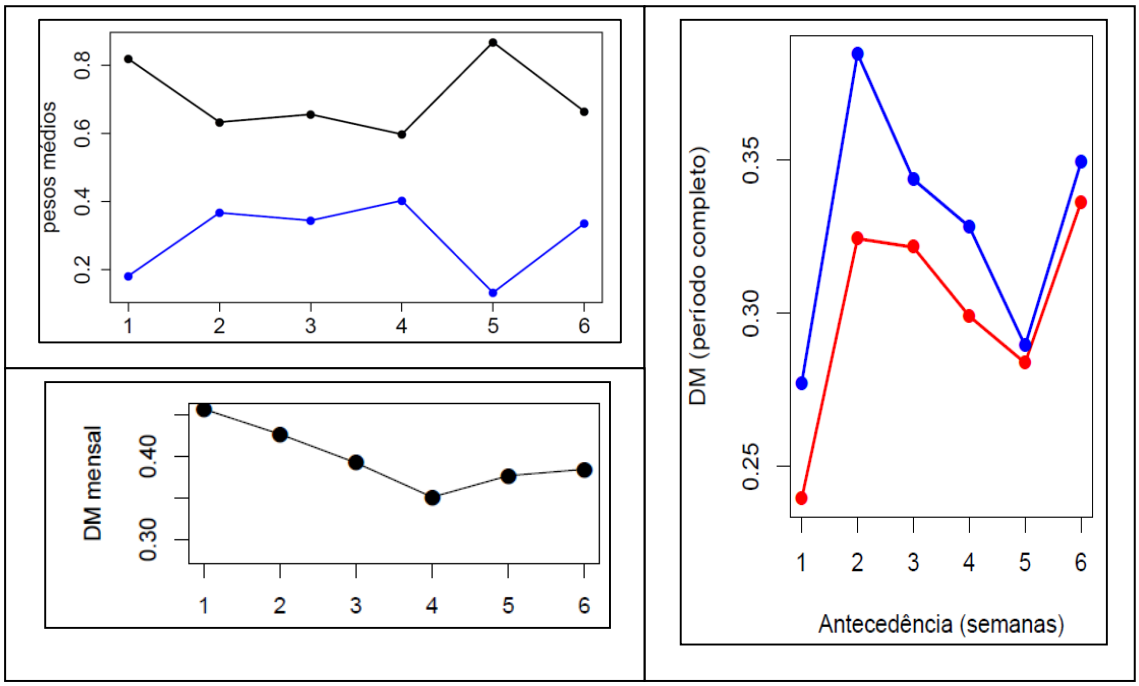

Figura D.8 - Assim como na Figura D.1, mas para a UHE Água Vermelha.

Bacia do rio Tietê

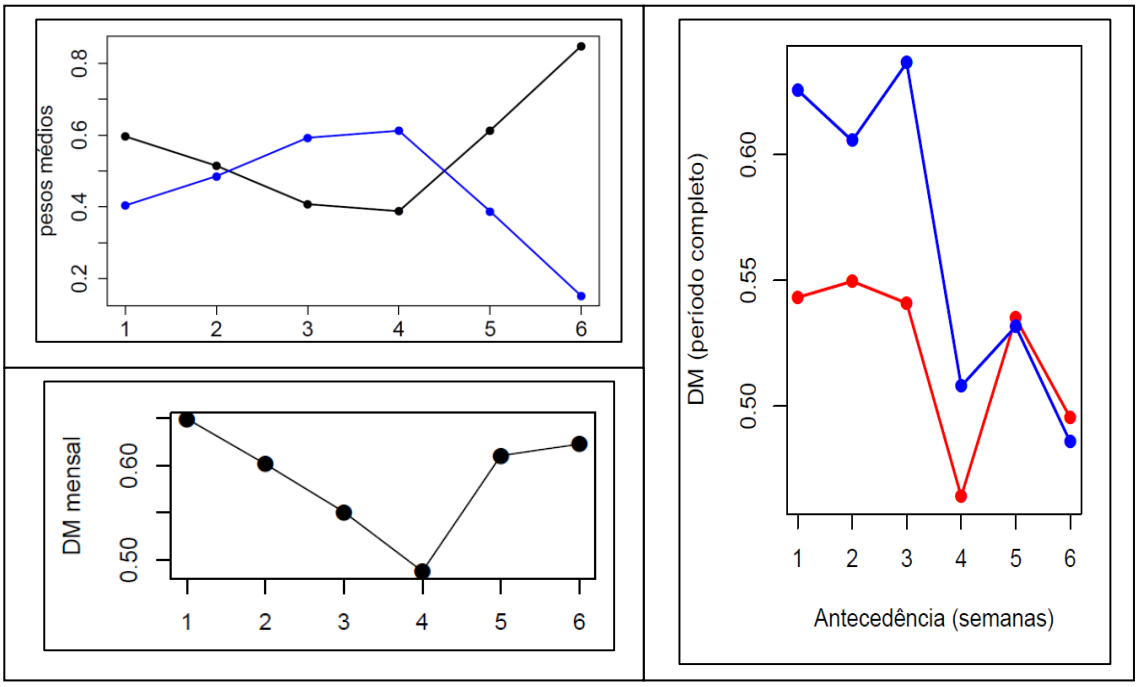

Figura D.9 - Assim como na Figura D.1, mas para a UHE Barra Bonita.

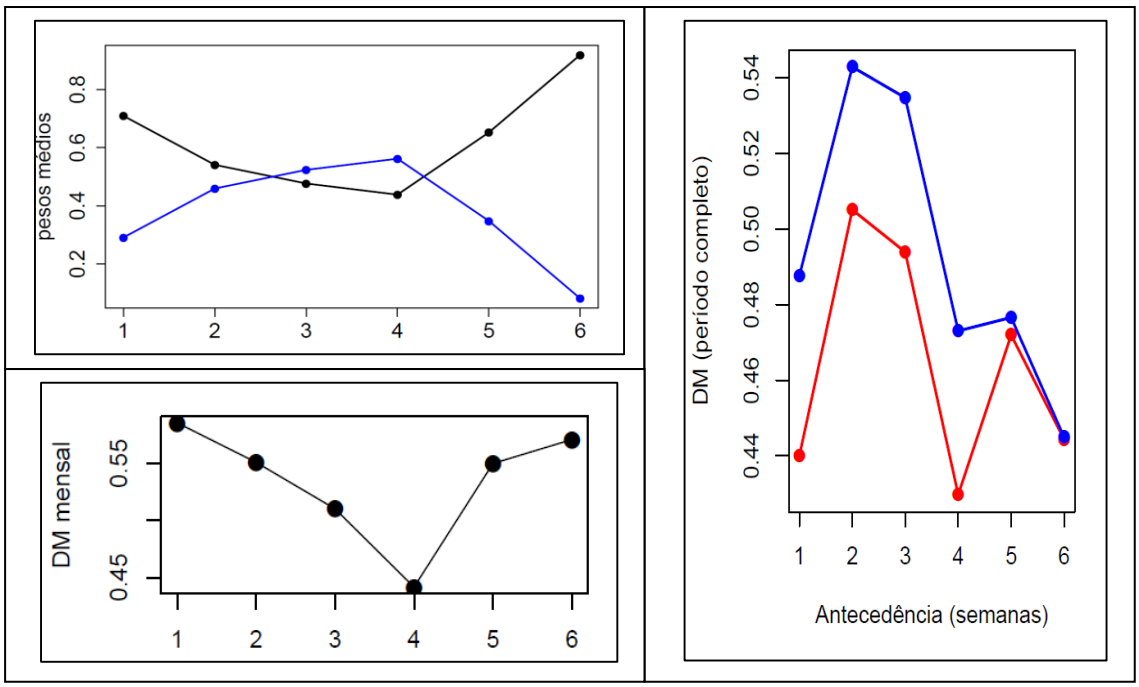

Figura D.10 - Assim como na Figura D.1, mas para a UHE Promissão.

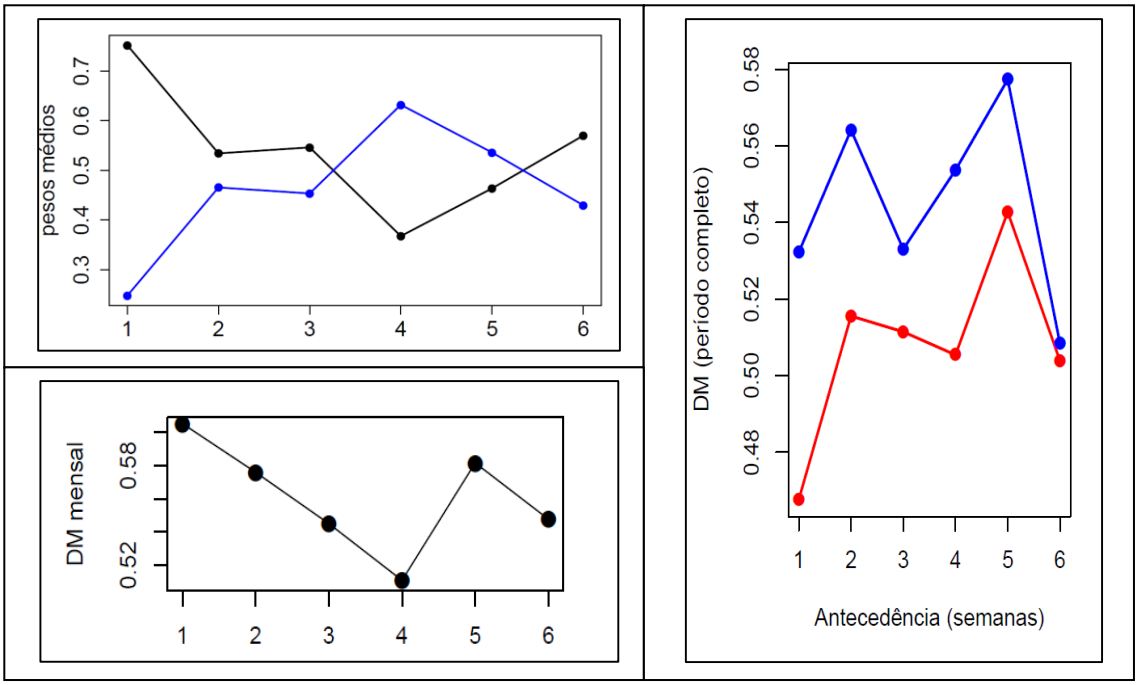

Figura D.11 - Assim como na Figura D.1, mas para a UHE Três Irmãos.

Bacia do rio Parapanema

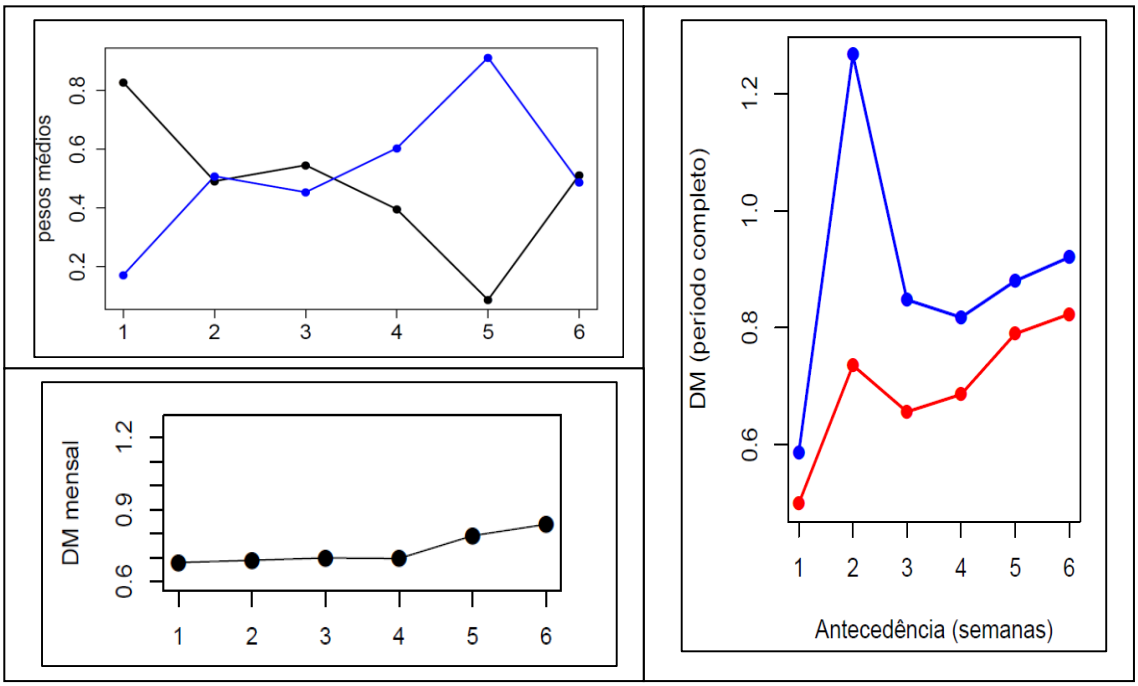

Figura D.12 - Assim como na Figura D.1, mas para a UHE Jurumirim (A. A. Laydner).

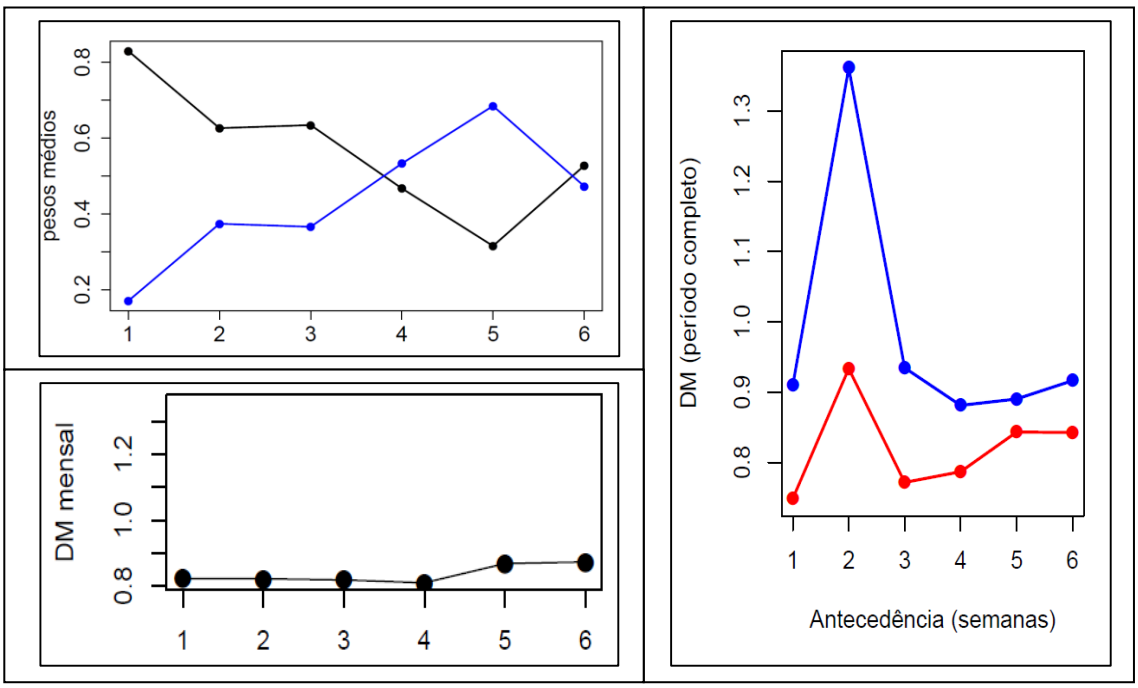

Figura D.13 - Assim como na Figura D.1, mas para a UHE Chavantes.

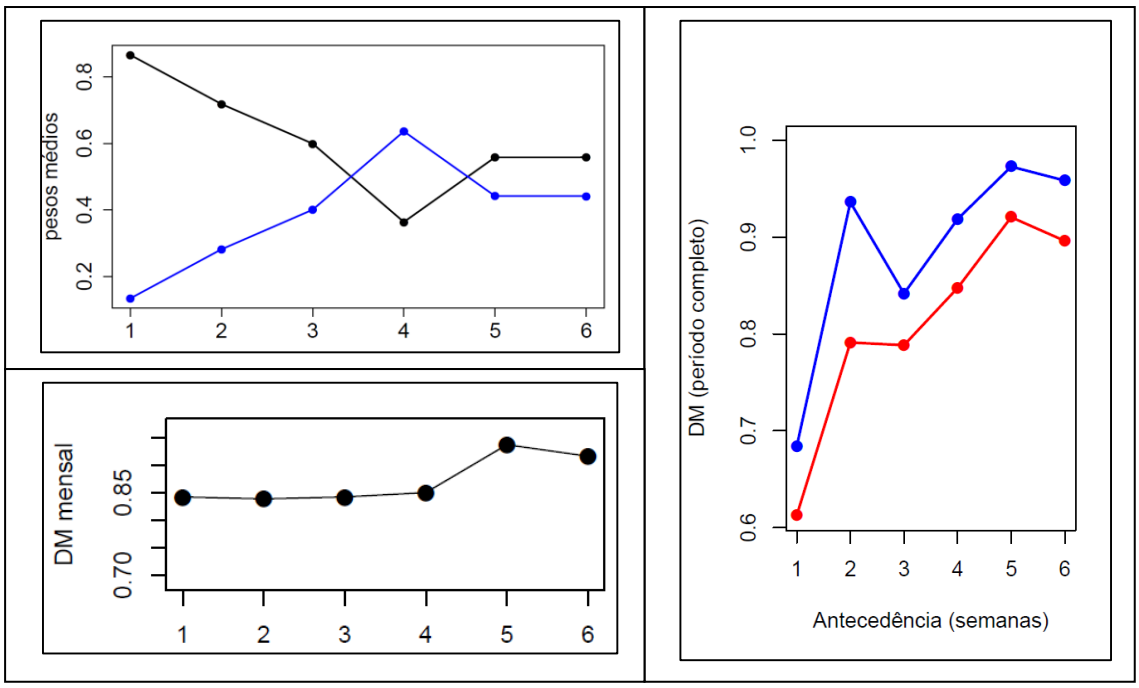

Figura D.14 - Assim como na Figura D.1, mas para a UHE Capivara.

Bacia do rio Paraná

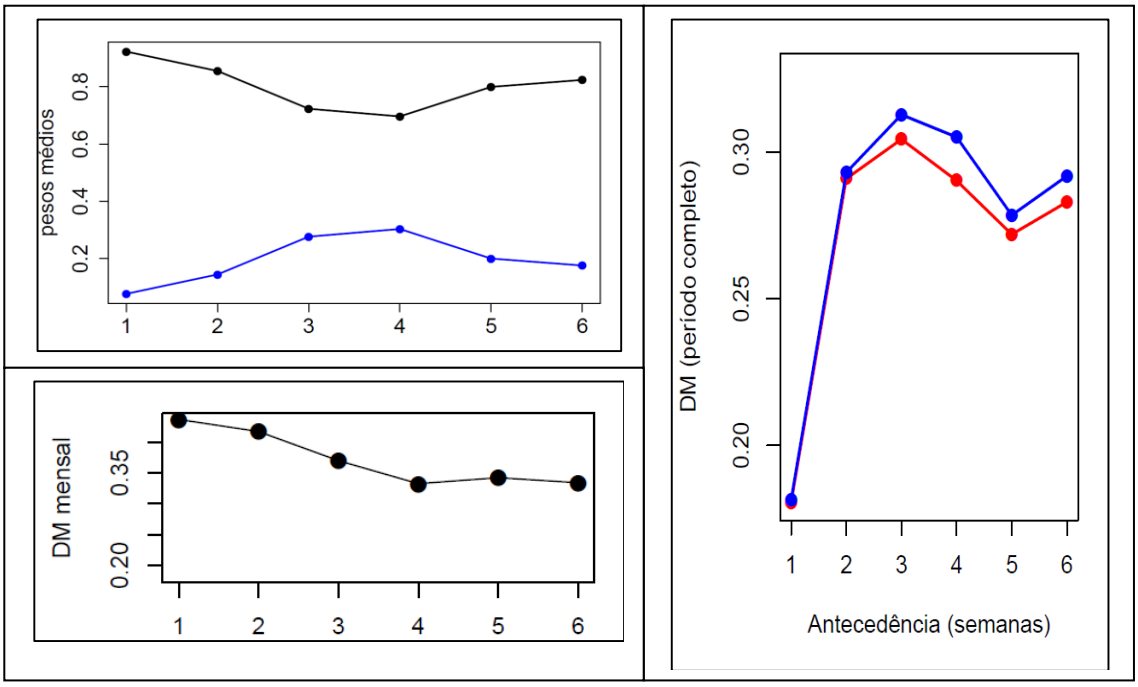

Figura D.15 - Assim como na Figura D.1, mas para a UHE Ilha Solteira.

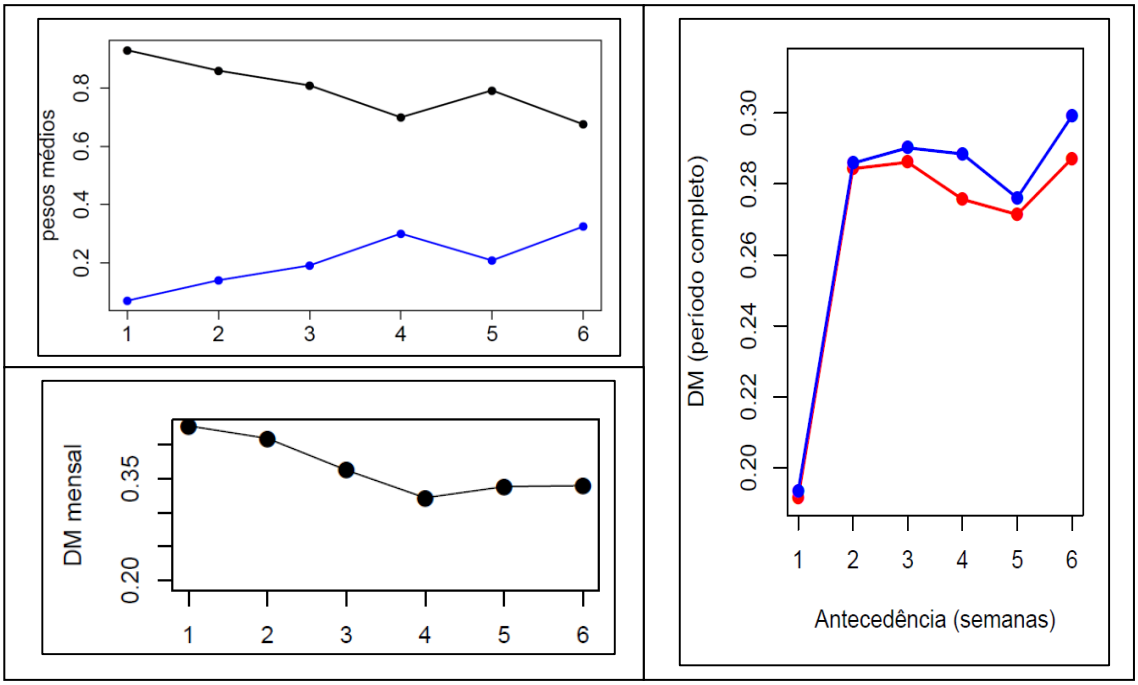

Figura D.16 - Assim como na Figura D.1, mas para a UHE Jupiá.

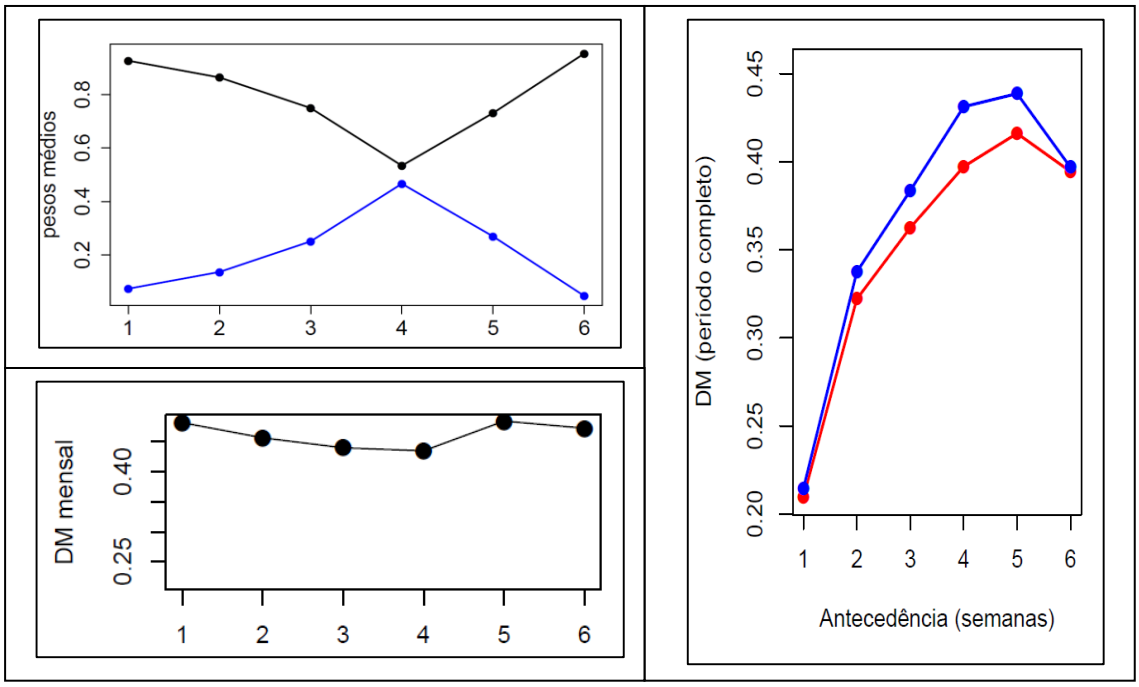

Figura D.17 - Assim como na Figura D.1, mas para a UHE Itaipu.

Bacia do rio Iguaçu

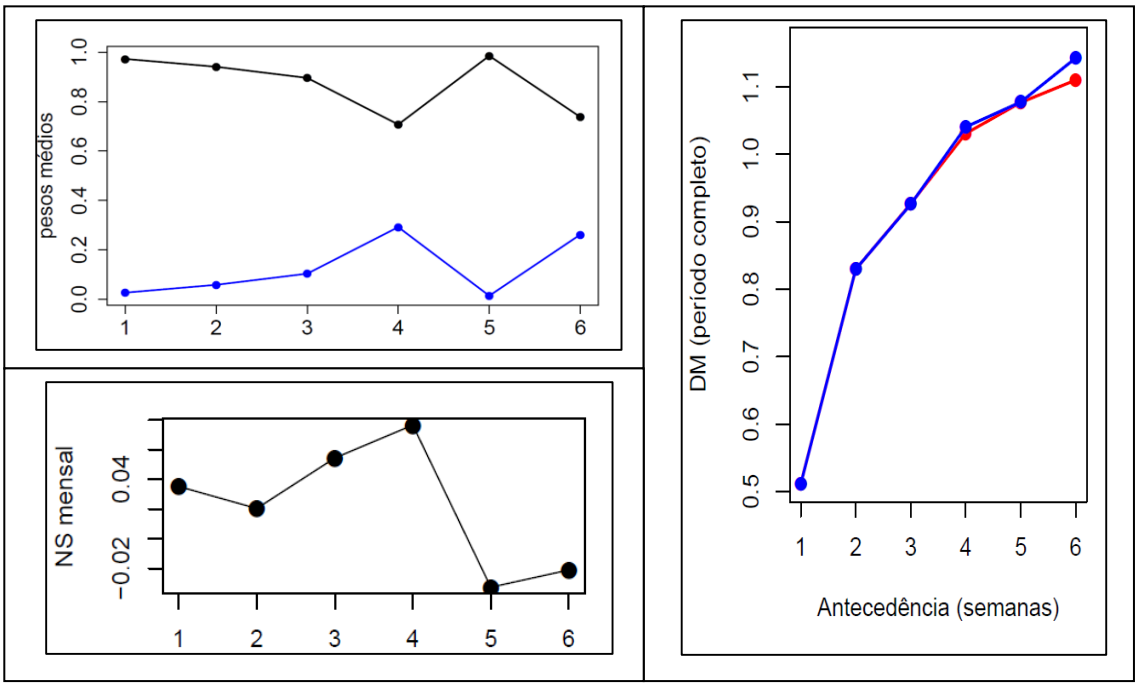

Figura D.18 - Resultados de uma a seis semanas de antecedência para a UHE Foz do Areia (G. B. Munhoz) no período entre 2009 e 2014: o painel superior da esquerda mostra a ponderação dos pesos dos modelos semanais (curva preta) e mensais (curva azul); o painel inferior da esquerda, o indicador NS que compara as previsões mensais (na escala semanal) e os valores semanais verificados. Por fim, o painel da direita traz o indicador NS para as previsões semanais realizadas pelo ONS (curva na cor azul) e para as previsões BMA (curva vermelha).

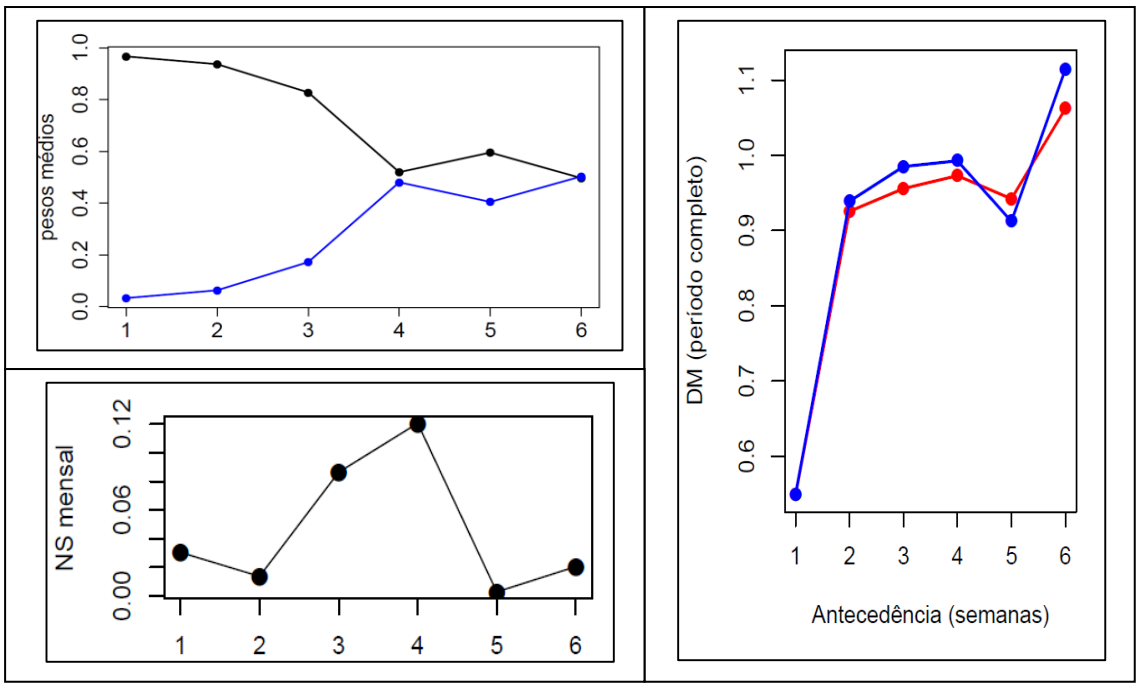

Figura D.19 - Assim como na Figura D.18, mas para a UHE Salto Santiago.

Bacia dos rios Uruguai e Jacuí

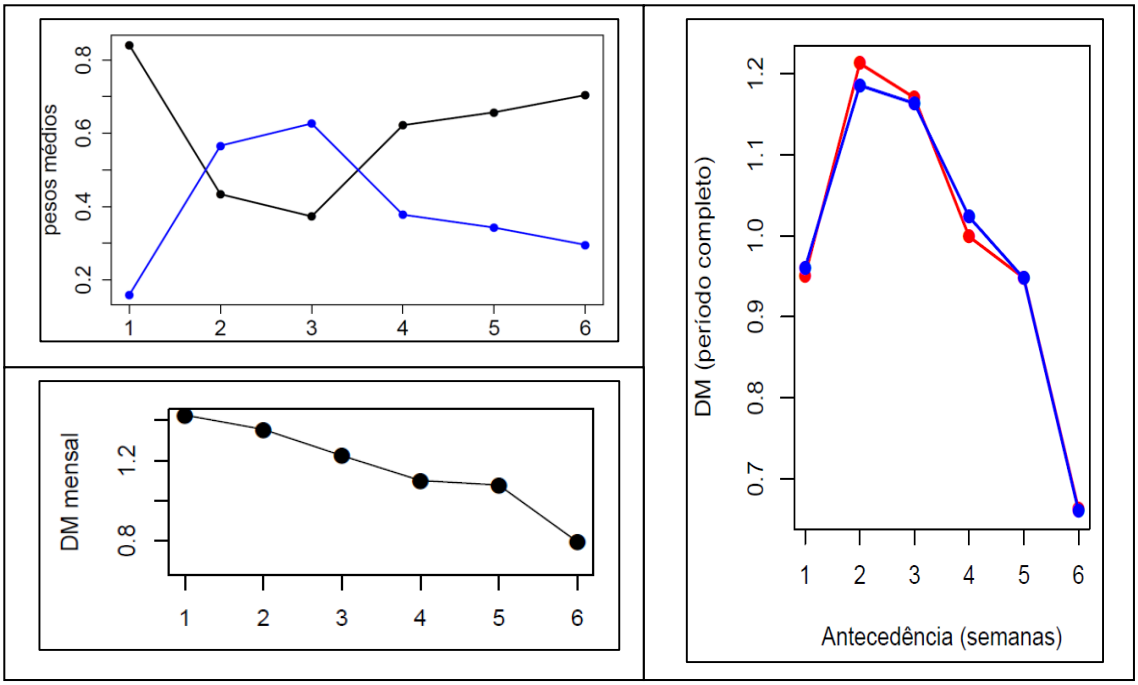

Figura D.20 - Assim como na Figura D.1, mas para a UHE Passo Fundo.

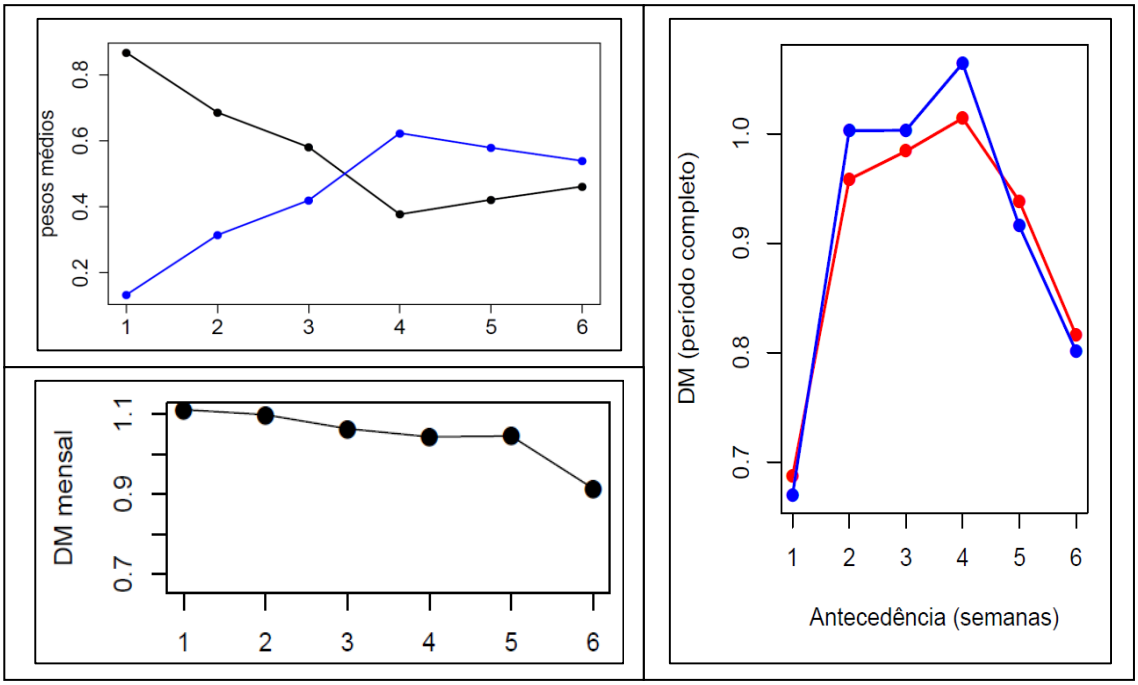

Figura D.21 - Assim como na Figura D.1, mas para a UHE Machadinho.

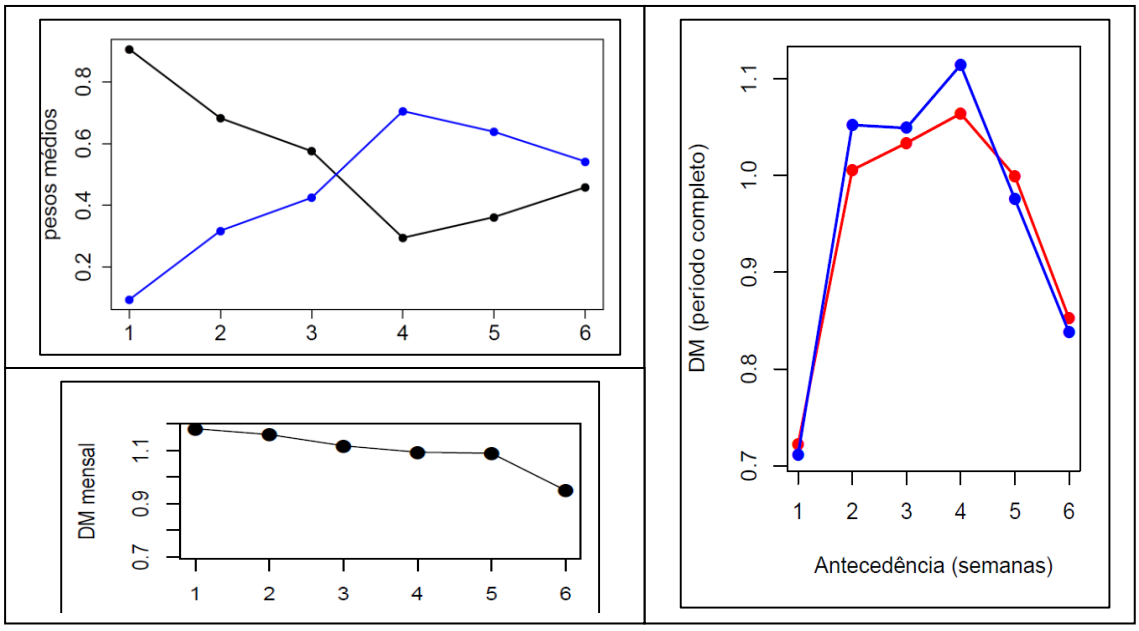

Figura D.22 - Assim como na Figura D.1, mas para a UHE Itá.

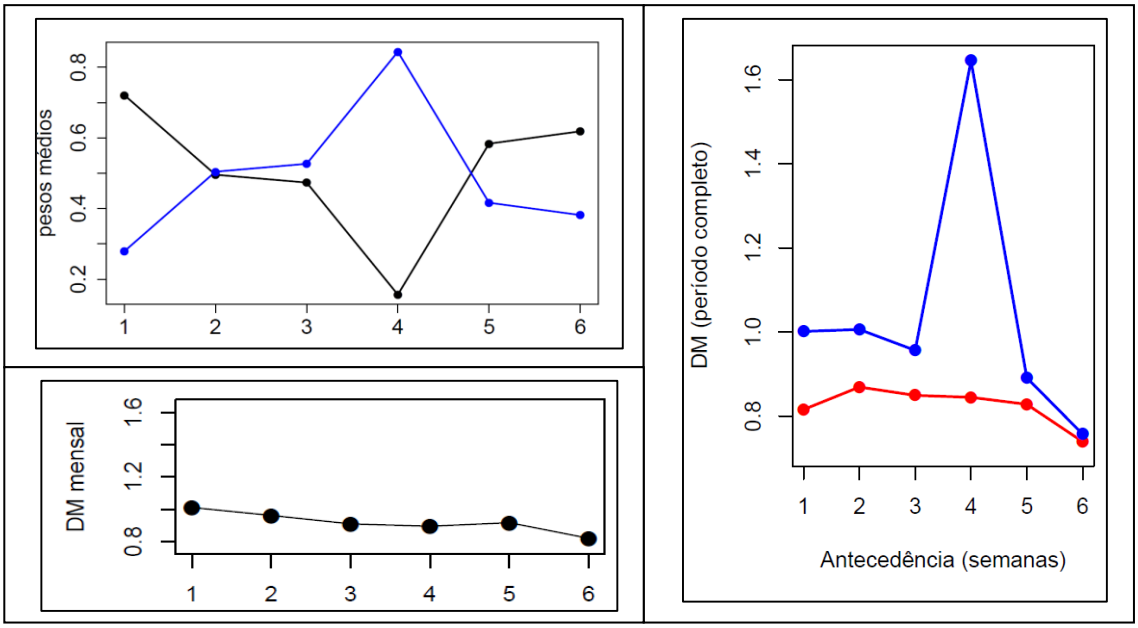

Figura D.23 - Assim como na Figura D.1, mas para a UHE Passo Real.

Bacias dos rios Sobradinho e Tocantins

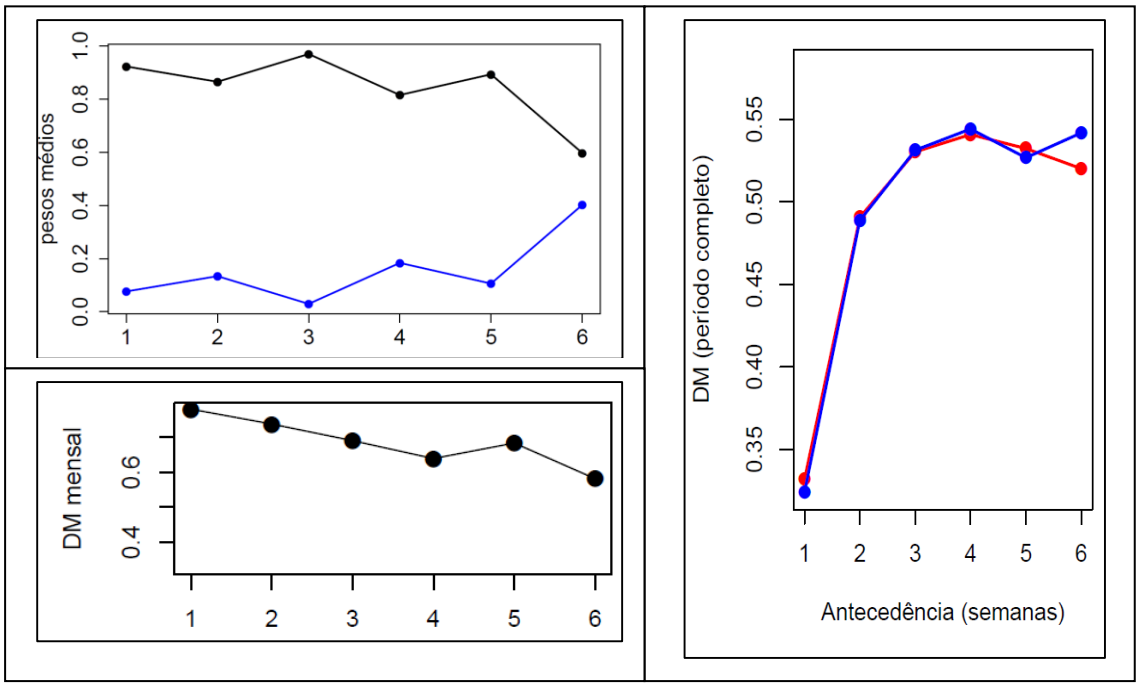

Figura D.24 - Assim como na Figura D.1, mas para a UHE Três Marias.

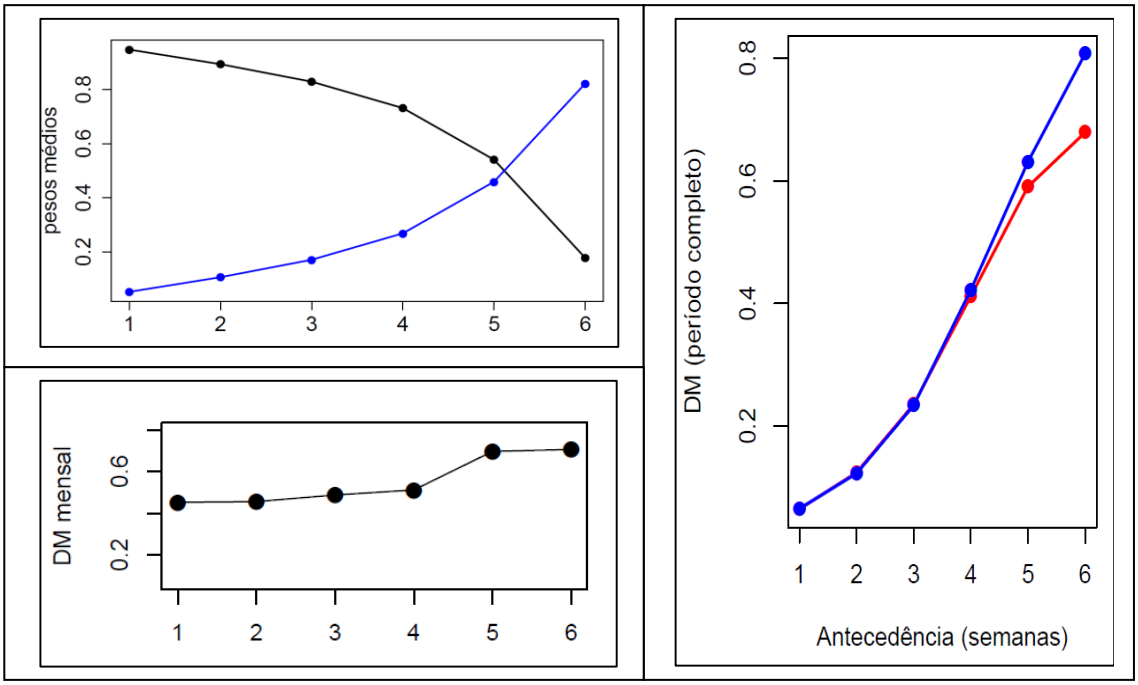

Figura D.25 - Assim como na Figura D.1, mas para a UHE Sobradinho.

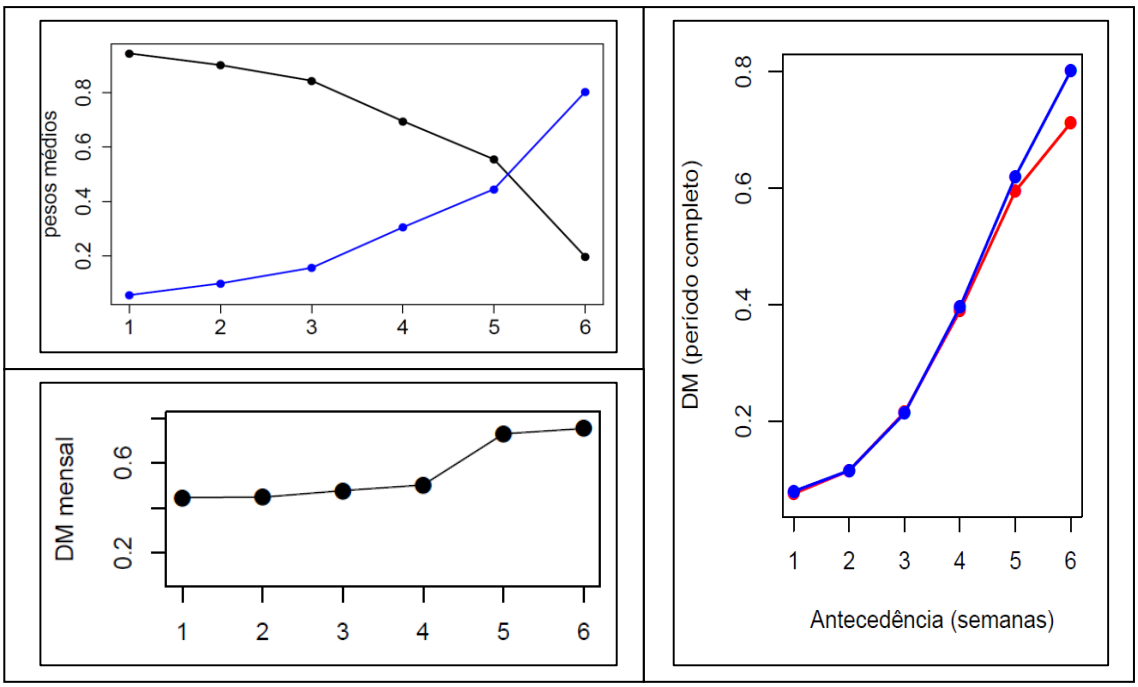

Figura D.26 - Assim como na Figura D.1, mas para a UHE Itaparica.

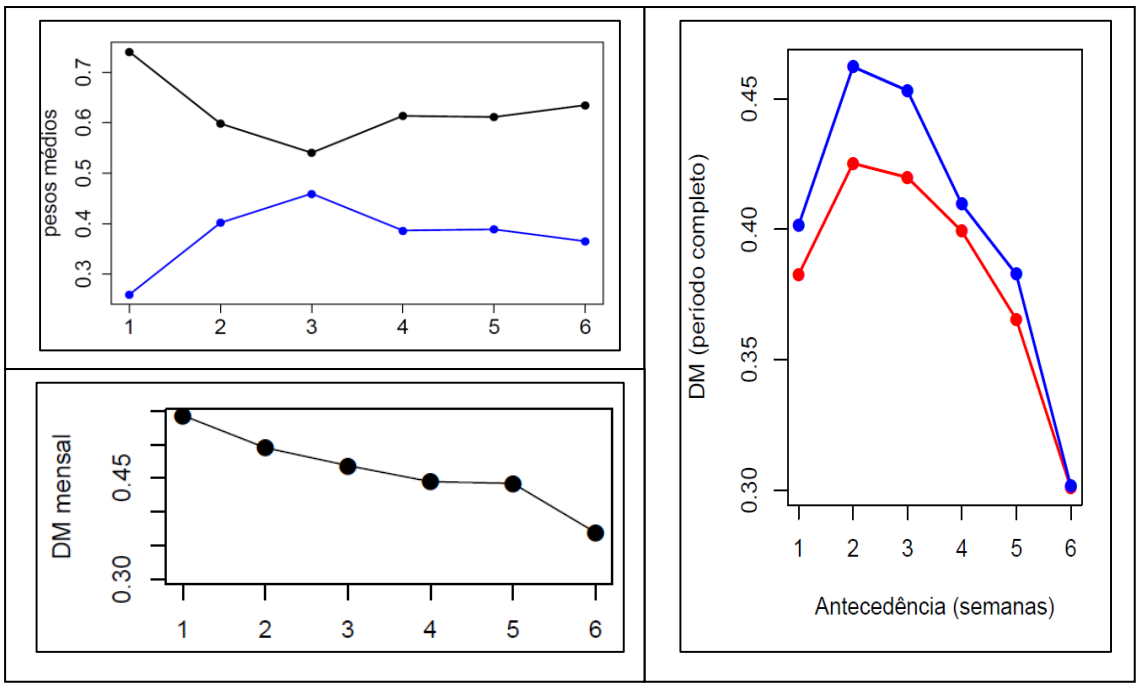

Figura D.27 - Assim como na Figura D.1, mas para a UHE Serra da Mesa.

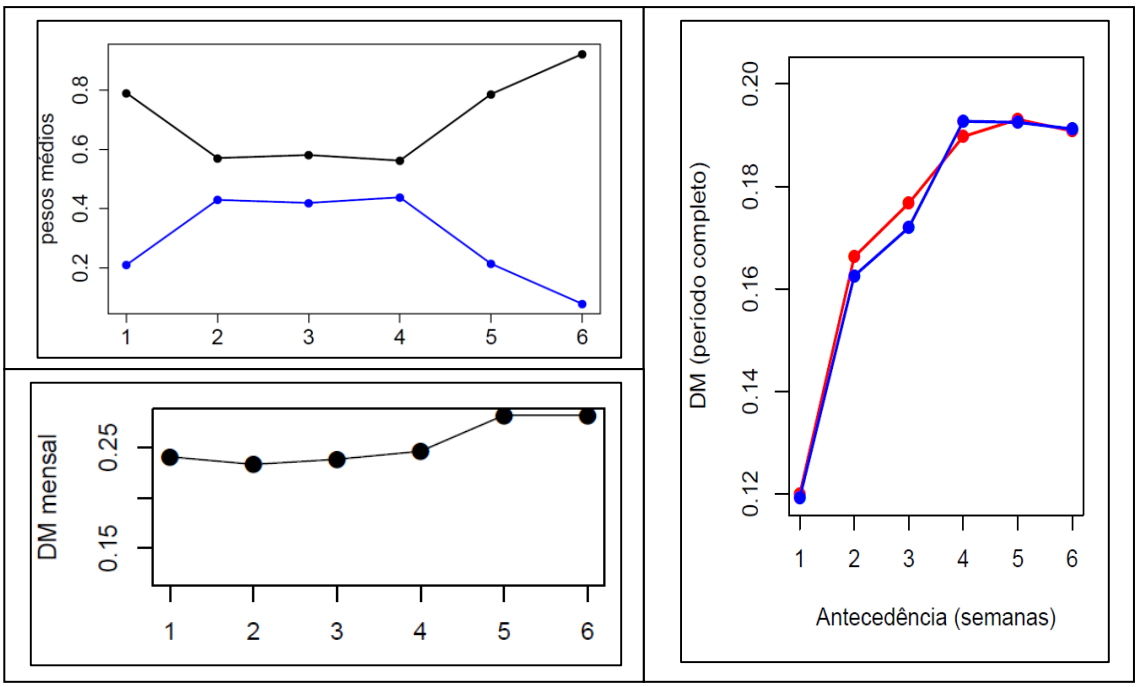

Figura D.28 - Assim como na Figura D.1, mas para a UHE Tucuruí.

APÊNDICE E - INDICADORES DE QUALIDADE MÉDIOS (COMPARATIVO DOS RESULTADOS BMA EM RELAÇÃO AO DESEMPENHO DOS MODELOS UTILIZADOS PELO ONS ENTRE JAN-2009 E SET-2014).

RMSE	1sem_anteced.	2sem_anteced.	3sem_anteced.	4sem_anteced.	5sem_anteced.	6sem_anteced.
FURNAS	12,61%	18,25%	5,28%	9,21%	0,09%	1,60%
MASCARENHAS	15,65%	18,46%	5,86%	9,56%	0,44%	1,74%
MARIMBONDO	10,87%	11,75%	4,31%	6,92%	0,93%	0,88%
AGUA.VERMELHA	10,75%	12,92%	4,69%	5,75%	0,58%	2,58%
EMBORCACAO	4,11%	4,49%	5,36%	6,69%	3,64%	12,37%
NOVA.PONTE	5,69%	6,37%	11,73%	9,11%	34,03%	0,27%
ITUMBIARA	2,95%	2,41%	2,70%	4,39%	2,07%	7,42%
SAO.SIMAO	-0,57%	0,62%	1,83%	2,56%	0,19%	0,90%
ILHA.SOLTEIRA	0,80%	0,77%	1,90%	3,21%	1,05%	1,45%
A.A.LAYDNER	10,01%	35,05%	15,96%	10,73%	7,54%	6,28%
CHAVANTES	11,53%	22,99%	11,57%	6,63%	3,60%	4,72%
CAPIVARA	6,26%	9,70%	3,71%	4,68%	3,06%	3,74%
GB.MUNHOZ	0,22%	0,31%	0,06%	0,96%	-0,01%	1,37%
SALTO.SANTIAGO	-0,03%	0,99%	2,07%	1,25%	-2,92%	2,52%
ITA	0,10%	5,13%	2,55%	6,15%	-0,55%	-0,59%
PASSO.FUNDO	5,76%	9,04%	4,44%	3,35%	0,36%	1,07%
PASSO.REAL	16,70%	12,40%	10,42%	54,44%	5,96%	3,58%
TRES.MARIAS	1,01%	1,25%	0,60%	0,03%	-0,58%	2,87%
SOBRADINHO	3,46%	4,50%	3,40%	3,18%	6,21%	14,45%
ITAPARICA	4,23%	5,81%	3,62%	3,40%	4,88%	11,07%
MACHADINHO	-0,16%	4,54%	2,55%	5,28%	-1,11%	-0,98%
BARRA.BONITA	10,35%	7,25%	11,85%	7,02%	0,05%	-0,80%
PROMISSAO	6,06%	4,68%	5,20%	6,77%	0,53%	-0,08%
TRES IRMAOS	8,14%	5,66%	2,64%	5,27%	3,72%	0,48%
JUPIA	0,85%	0,51%	0,89%	2,64%	0,67%	1,76%
ITAIPU	1,52%	2,78%	3,15%	4,81%	2,94%	0,33%
SERRA.MESA	3,86%	7,24%	6,87%	2,99%	6,21%	4,38%
TUCURUI	5,18%	5,02%	5,53%	10,93%	3,90%	0,59%
valores médios	**5,64%**	**7,89%**	**5,03%**	**7,07%**	**3,12%**	**3,07%**

Tabela E.1 – Resultados do indicador RMSE para as previsões com uma a seis semanas de antecedência. Valores positivos indicam ganho com as ponderações BMA.

MAPE	1sem_anteced.	2sem_anteced.	3sem_anteced.	4sem_anteced.	5sem_anteced.	6sem_anteced.
FURNAS	2,83%	1,00%	0,93%	3,97%	0,54%	1,81%
MASCARENHAS	0,70%	0,67%	1,37%	4,42%	0,63%	1,92%
MARIMBONDO	1,45%	2,73%	2,74%	6,87%	1,00%	0,85%
AGUA.VERMELHA	3,91%	3,16%	2,56%	7,35%	2,83%	1,62%
EMBORCACAO	-0,48%	-2,73%	-3,98%	-5,48%	-7,18%	-5,57%
NOVA.PONTE	-1,20%	-2,53%	-0,47%	-0,40%	0,10%	-0,22%
ITUMBIARA	-1,16%	-1,69%	-2,22%	-2,50%	-0,74%	0,36%
SAO.SIMAO	-0,11%	-1,41%	-0,25%	-0,38%	0,62%	2,33%
ILHA.SOLTEIRA	-0,75%	-1,00%	0,46%	2,57%	2,88%	3,51%
A.A.LAYDNER	2,98%	11,11%	2,93%	2,55%	-8,42%	3,36%
CHAVANTES	3,54%	7,41%	3,02%	3,12%	-3,34%	2,30%
CAPIVARA	2,82%	3,73%	1,51%	0,55%	1,33%	1,18%
GB.MUNHOZ	-0,62%	-1,01%	-0,53%	-1,51%	0,30%	3,74%
SALTO.SANTIAGO	0,17%	-0,07%	-0,03%	0,79%	4,48%	4,10%
ITA	-4,42%	-2,99%	-3,64%	-8,12%	-4,56%	-2,77%
PASSO.FUNDO	-5,09%	-16,79%	-7,15%	-1,13%	-1,04%	-2,74%
PASSO.REAL	0,36%	-1,30%	-3,36%	-3,70%	-3,61%	-9,96%
TRES.MARIAS	-4,06%	-2,28%	-0,59%	1,04%	-0,92%	1,65%
SOBRADINHO	-1,85%	-4,57%	-5,28%	-2,25%	-3,76%	-6,27%
ITAPARICA	3,45%	-3,31%	-5,00%	-3,96%	-5,98%	-10,44%
MACHADINHO	-6,23%	-1,96%	-3,10%	-6,03%	-2,83%	-1,33%
BARRA.BONITA	2,12%	-0,09%	1,53%	1,24%	-2,46%	-3,12%
PROMISSAO	6,85%	1,75%	1,31%	1,25%	0,71%	1,35%
TRES IRMAOS	4,81%	3,19%	0,54%	5,58%	1,77%	0,61%
JUPIA	-0,02%	-0,55%	0,58%	3,10%	2,55%	5,72%
ITAIPU	0,38%	1,42%	3,43%	2,53%	2,70%	0,37%
SERRA.MESA	0,45%	-1,71%	-1,82%	-2,22%	-4,43%	-6,58%
TUCURUI	-3,53%	-6,21%	-7,04%	-5,60%	-3,85%	-0,21%
valores médios	**0,26%**	**-0,57%**	**-0,77%**	**0,13%**	**-1,10%**	**-0,44%**

Tabela E.2 – Resultados do indicador MAPE para as previsões com uma a seis semanas de antecedência. Valores positivos indicam ganho com as ponderações BMA.

NS	1sem_anteced.	2sem_anteced.	3sem_anteced.	4sem_anteced.	5sem_anteced.	6sem_anteced.
FURNAS	-6,83%	-22,91%	-5,32%	-6,94%	-0,06%	-1,26%
MASCARENHAS	-8,30%	-21,23%	-5,57%	-6,95%	-0,31%	-1,34%
MARIMBONDO	-5,48%	-8,62%	-3,06%	-3,79%	-0,47%	-0,52%
AGUA.VERMELHA	-5,23%	-9,34%	-3,21%	-3,31%	-0,30%	-1,89%
EMBORCACAO	-4,94%	-6,25%	-9,12%	-12,61%	-6,30%	-14,52%
NOVA.PONTE	-6,75%	-9,34%	-16,49%	-12,66%	-51,86%	-0,34%
ITUMBIARA	-2,32%	-3,00%	-3,35%	-6,14%	-2,40%	-6,77%
SAO.SIMAO	0,18%	-0,50%	-1,71%	-2,32%	-0,14%	-0,56%
ILHA.SOLTEIRA	-0,25%	-0,49%	-1,28%	-1,97%	-0,57%	-0,86%
A.A.LAYDNER	-16,82%	-166,48%	-50,04%	-36,91%	-38,28%	-43,53%
CHAVANTES	-54,50%	-326,26%	-58,05%	-34,83%	-24,24%	-35,76%
CAPIVARA	-17,16%	-57,22%	-20,72%	-35,53%	-40,48%	-42,04%
GB.MUNHOZ	-0,25%	-1,21%	-0,37%	-12,72%	0,29%	257,29%
SALTO.SANTIAGO	0,05%	-7,72%	-18,57%	-12,05%	26,33%	-239,19%
ITA	-0,28%	-30,66%	-18,50%	-57,70%	4,13%	2,78%
PASSO.FUNDO	-14,87%	-51,82%	-24,04%	-13,33%	-1,72%	-1,80%
PASSO.REAL	-60,71%	-52,61%	-42,61%	-263,11%	-27,13%	-11,81%
TRES.MARIAS	-0,42%	-1,10%	-0,64%	-0,04%	0,72%	-3,91%
SOBRADINHO	-0,26%	-0,58%	-1,12%	-2,80%	-11,19%	-36,79%
ITAPARICA	-0,40%	-0,68%	-1,02%	-2,63%	-8,76%	-31,16%
MACHADINHO	0,39%	-24,69%	-16,55%	-45,11%	6,97%	4,47%
BARRA.BONITA	-16,85%	-12,18%	-19,73%	-8,19%	-0,08%	1,21%
PROMISSAO	-7,28%	-7,23%	-7,84%	-7,69%	-0,75%	0,10%
TRES IRMAOS	-11,29%	-9,17%	-4,31%	-8,45%	-7,04%	-0,81%
JUPIA	-0,30%	-0,32%	-0,56%	-1,57%	-0,39%	-1,14%
ITAIPU	-0,67%	-2,27%	-3,07%	-5,53%	-3,65%	-0,36%
SERRA.MESA	-3,30%	-7,42%	-6,76%	-2,62%	-4,53%	-2,17%
TUCURUI	-0,59%	-0,88%	-1,03%	-2,28%	-0,92%	-0,14%
valores médios	**-8,77%**	**-30,08%**	**-12,31%**	**-21,78%**	**-6,90%**	**-7,60%**

Tabela E.3 – Resultados do indicador NS para as previsões com uma a seis semanas de antecedência. Valores negativos indicam ganho com as ponderações BMA.

DM	1sem_anteced.	2sem_anteced.	3sem_anteced.	4sem_anteced.	5sem_anteced.	6sem_anteced.
FURNAS	17,31%	28,00%	7,21%	11,66%	0,37%	2,57%
MASCARENHAS	20,83%	27,56%	7,96%	12,24%	0,75%	2,77%
MARIMBONDO	14,55%	16,21%	6,26%	10,50%	1,40%	1,33%
AGUA.VERMELHA	15,77%	18,61%	6,88%	9,76%	1,98%	3,88%
EMBORCACAO	6,32%	5,91%	7,30%	9,00%	3,72%	15,81%
NOVA.PONTE	8,55%	8,95%	17,68%	13,09%	58,29%	0,31%
ITUMBIARA	3,90%	3,20%	3,59%	6,27%	2,87%	10,28%
SAO.SIMAO	-0,66%	0,51%	2,75%	3,83%	0,44%	1,98%
ILHA.SOLTEIRA	0,60%	0,73%	2,75%	5,12%	2,40%	3,13%
A.A.LAYDNER	17,49%	72,16%	29,19%	19,20%	11,44%	11,82%
CHAVANTES	21,51%	45,81%	21,11%	11,99%	5,58%	8,75%
CAPIVARA	11,63%	18,31%	6,76%	8,39%	5,67%	6,95%
GB.MUNHOZ	-0,10%	0,03%	-0,09%	0,95%	0,06%	2,94%
SALTO.SANTIAGO	0,02%	1,45%	3,06%	2,04%	-3,18%	4,96%
ITA	-1,50%	4,63%	1,49%	4,73%	-2,37%	-1,69%
PASSO.FUNDO	1,03%	-2,24%	-0,59%	2,43%	-0,06%	-0,40%
PASSO.REAL	22,78%	15,75%	12,60%	94,72%	7,62%	2,41%
TRES.MARIAS	-2,39%	-0,40%	0,19%	0,62%	-1,05%	4,16%
SOBRADINHO	0,99%	-1,31%	-0,79%	2,54%	6,88%	18,91%
ITAPARICA	5,24%	0,25%	-0,90%	1,65%	4,15%	12,52%
MACHADINHO	-2,52%	4,65%	1,94%	4,95%	-2,42%	-1,79%
BARRA.BONITA	15,18%	10,17%	17,67%	9,45%	-0,62%	-1,94%
PROMISSAO	10,83%	7,48%	8,27%	10,09%	0,98%	0,13%
TRES IRMAOS	13,81%	9,42%	4,22%	9,55%	6,37%	0,90%
JUPIA	1,06%	0,55%	1,42%	4,64%	1,73%	4,18%
ITAIPU	2,34%	4,79%	5,81%	8,57%	5,43%	0,61%
SERRA.MESA	4,95%	8,71%	7,92%	2,58%	4,72%	0,26%
TUCURUI	-0,60%	-2,23%	-2,69%	1,58%	-0,30%	0,21%
valores médios	**7,46%**	**10,99%**	**6,39%**	**10,08%**	**4,39%**	**4,14%**

Tabela E.4 – Resultados do indicador DM para as previsões com uma a seis semanas de antecedência. Valores positivos indicam ganho com as ponderações BMA.